







RESULT 4  
 AX780857 AX780857 2788 bp DNA linear PAT 14-JUL-2003  
 LOCUS DEFINITION Sequence 3014 from Patent WO03039443.  
 ACCESSION AX780857  
 VERSION GI:32697881  
 KEYWORDS SOURCE Homo sapiens (human)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homo.  
 REFERENCE AUTHORS Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,  
 Dugas,M., Eisls,R., Brors,B. and Mergenthaler,S.  
 TITLE Patents  
 JOURNAL Patent: WO 03039443-A 3014-15-NAY-2003;  
 Deutsches Krebsforschungszentrum (DE) ;  
 Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,  
 PD Dr. Dr. Schoch (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)  
 FEATURES Location/Qualifiers  
 1..2788  
 /organism="Homo sapiens"  
 /mol type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match Similarity 97.4%; Score 1152.4; DB 6; Length 2788;  
 Best Local Similarity 98.4%; Pred No. 1..8e-297;  
 Matches 1153; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy	12	CAAGGCCCCACGCCCTGTGTCCTGTCAGAAGATGCCTAACGGCCCTGGCCCTCCCTCC	71
Db	1	CAAGGCCCCACGCCCTGTGTCAGAAGATGCCTAACGGCCCTGGCCCTCCCTCC	60
Qy	72	TGCTCTGGCTTGCTGGTGGAGGGTCCCGACTCCAGATGGGGATGGGAGC	131
Db	61	TGCTCTGGCTTGCTGGTGGAGGGTCCCGACTCCAGATGGGAGC	120
Qy	132	TGATCCATTCTGGTACAAACTGGTCACTGGCTGGAGATGGGGATGGGAGC	191
Db	121	TGATCCATTCTGGTACAAACTGGTCACTGGCTGGAGATGGGAGC	180
Qy	192	CACCTAGGCTCTCCCTGAGATCTCCCTGGGTGAGAGATCTGGGTGCTAGGACCA	251
Db	181	CACCTAGGCTCTCCCTGAGATCTCCCTGGGTGAGAGATCTGGGTGCTAGGACCA	240
Qy	252	AGGACACTGGCAGACTTCAGAGGGCCCCAAAGGCCCTAACCTGTCAGGCAGGCTAG	311
Db	241	AGGACACTGGCAGACTTCAGAGGGCCCCAAAGGCCCTAACCTGTCAGGCAGGCTAG	300

RESULT 5  
 BC042041 LOCUS BC042041 mRNA linear PRI 07-OCT-2003  
 DEFINITION Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (cDNA  
 clone MGIC:49845 IMAGE:4429896), complete cds.  
 ACCESSION BC042041  
 VERSION BC042041.1 GI:27459842  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2538)

AUTHORS Strauberg,R.L., Peingold,E.A., Grouse,I.H., Derge,J.G., Schuller,G.D., Klausner,R.D., Collins,F.S., Wagner,L., Sheeney,C.M., Bhat,N.K., Altschul,S.F., Zeeberg,B., Bustow,K.H., Schaefer,C.P., Bhat,N.K., Hopkinson,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hecht,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Staszewski,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiki,Yuki,S., Carninci,P., Prange,C., Raha,S.S., Loqueland,N.A., Peters,G.J., Abramson,R.D., Mulligan,S.J., Bosak,S.A., McIwan,P.J., McKernan,K.J., Malek,J.A., Gunnarsson,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,J.J., Hulyk,S.W., Vaillancourt,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmitz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smialius,D.E., Schnarch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBLISHER 1247732

PUBLISHED 1247732

(bases 1 to 2538)

AUTHORS Strauberg,R.

TITLE Direct Submission

JOURNAL Submitted (23-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

EMAIL Email: cgapbs-2@mail.nih.gov

Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (IIML)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-sgic.stanford.edu>

Contact: (Dickson, Mark) mcdopaxil.stanford.edu

Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

clone distribution: MGC clone distribution information can be found through the J.M.A.G.E. Consortium/LiML at: <http://image.lnl.gov>  
series: IRAK Plate: 88 Row: a Column: 20  
this clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28416422.

SOURCES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_id="MGC:49845 IMAGE:4429986"
/tissue_type="Prostate, adenocarcinoma"
/clone_id="NIH MGC_91"
/lab_host="DH10E"
/note=vector: pCMV-SPORT6"
1. 2538
/gene="SLA2"
/note=synonyms: FLJ21992, SLAP-2, C20orf156, MGCG49845,
SLAP2
/db_xref="LocusID:84174"
/db_xref="MIM:606577"
363_1148
/codon_start=1
/product="src-like-adaptor 2, isoform a"
/protein_id="3AHA2041_1"
/db_xref="GI:27469843"
/db_xref="LocusID:84174"
/rtranslation="MCASTGPKPDKDPEKQEVQCDYMMRDPDKTMYAVIICSD
CDS
gene

```

Db 781 CATCTGGACCCGATCAGAACACTAACAGATTGCAATGGCTGGCTGTACA 840  
 Qy 876 TCTCACCGGGCTTACCTTCCTCACTCCAGGCCCTGGGACCTTAATCTGAGCTGG 935  
 Db 841 TCTCACCGGGCTTACCTTCCTCACTAACGGCCCTGGGACCTTAATCTGAGCTGG 900  
 Qy 936 CGGTAGACATCTGTGCTACTAAGGCCCTGTGCTGGAGGCTGCCCCGCTCC 995  
 Db 901 CGGTAGACATCTGTGCTACTAAGGCCCTGTGCTGGAGGCTGCCCCGCTCC 960  
 Qy 996 CTGGAAAGGTTAACCCCTAACCTGTGAGGAGACCAACTAACCTGAAAGAGC 1055  
 Db 961 CTGGAAAGGTTAACCCCTAACCTGTGAGGAGACCAACTAACCTGAAAGAGC 1020  
 Qy 1056 TGGAAAGGTTAACCCCTAACCTGTGAGGAGACCAACTAACCTGAAAGAGC 1115  
 Db 1021 TGGCAAGCTCCCTCTGGTCTGGTCTGGTCTGGTCTGGTCTGGTCTGG 1080  
 Qy 1116 GTCTCGGGAGTCCTCACTACATAGCCCTGAATGAGGGCTCTCTTGTGATG 1175  
 Db 1081 GTCTCGGGAGTCCTCACTACATAGCCCTGAATGAGGGCTCTCTTGTGATG 1140  
 Qy 1176 ATGCCTAG 1183  
 Db 1141 ATGCCTAG 1148

RESULT 6  
 AK025645 AK025645 2415 bp mRNA linear PRI 13-SEP-2003  
 DEFINITION Homo sapiens cDNA: FLJ21992 fis clone HBPP06554.  
 ACCESSION AK025645  
 VERSION AK025645.1 GI:10438227  
 KEYWORDS Oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikeda,Y.,  
 Okamoto,S.,'Oitaka,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,  
 Shiba,bara,T., Tanaka,T., Nakamura,T., Isogai,T., and Sugano,S.  
 TITLE NEDO human cDNA sequencing project  
 JOURNAL unpublished  
 REFERENCE 2 (bases 1 to 2415)  
 AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,  
 Shiba,bara,T., Tanaka,T., and Nakamura,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-AUG-2000) Sunio Sugano, Institute of Medical Science,  
 University of Tokyo, Laboratory of Genome Structure Analysis, Human  
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,  
 Japan (E-mail: flicona@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,  
 Fax: 81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan, cDNA full insert  
 Sequencing: Research Association for Biotechnology, cDNA library  
 construction, 5' - & 3' end one pass sequencing: Department of  
 Virology and Human Genome Center, Institute of Medical Science,  
 University of Tokyo (partly supported by Science and Technology  
 Agency).

FEATURES source  
 1..2415  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GI:10438228"  
 /comment="HEP06554"  
 /cell\_line="HEPG2"  
 /cell\_type="hepatoma"  
 /clone\_id="HEP"  
 /note="cloning vector pME18SPL3"  
 /note="unnamed protein product"

CDS 61..846  
 /note="unnamed protein product"

LOCUS	AX511153	DEFINITION	Sequence 4 from Patent WO0242452.	786 bp	DNA	linear	PAT 27-SEP-2002	Qy	1118	CTCCGGAGTCCCTAGCTCTACATCAGCTGATGAACTGGGTCTCTTGGAT 1177
DEFINITION	Sequence 4 from Patent WO0242452.							Db	721	CTCCGGAGTCCCTAGCTCTACATCAGCTGATGAACTGGGTCTCTTGGAT 780
VERSION	AX511153	KEYWORDS								
SOURCE	Homo sapiens (human)	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE	McGlade,J.C. and Loretto,M.P.	AUTHORS								
TITLE	Adapter gene	JOURNAL	Patent: WO 0242452-A 4 30-MAY-2002; The Hospital for Sick Children (CA) Location/Qualifiers							
FEATURES	1 .786 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	ORIGIN	Query Match 66.3%; Score 784.4; DB 6; Length 786; Best Local Similarity 99.9%; Pred. No. 6.9e-199; Matches 785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Qy	398 ATGGGAAGTTGCCCTAGGAGAAGAAAATCTGCAAGGCCAACGCTTCTAGTCCTCTCGTC 457	Db	1 ATGGGAAGTTGCCCTAGGAGAAGAAAATCTGCAAGGCCAACGCTTCTAGTCCTCTCGTC 60							
Qy	458 CAAGGCAGGGACCTGTGAACCATGGAAGGAGAACGGCAAGGCCGGCCCTG 517	Db	61 CAAGGCAGGGACCTGTGACCATGGAAGGAGAACGGCAAGGCCGGCCCTG 120							
Qy	518 GGCAAGTTCCGGAGGTGCCGGCCGGAGCTGGTGAGACTGGGGAGCCATTGACC 577	Db	121 GGCAAGTTCCGGAGGTGCCGGCCGGAGCTGGTGAGACTGGGGAGCCATTGACC 180							
Qy	578 ATTCGCTCTCTGAGGATGAGGAGACTGGGAGCTGCTGTAACCTCTCAGGAGAGAT 637	Db	181 ATTCGCTCTCTGAGGATGAGGAGACTGGGAGCTGCTGTAACCTCTCAGGAGAGAT 240							
Qy	638 AACATCCCCGGGCAACGTGGCAAAGGCTCCATGGGTGGCTGTATGAGGCCCTGAGC 697	Db	241 AACATCCCCGGGCAACGTGGCTCCATGGGTGGCTGTATGAGGCCCTGAGC 300							
Qy	698 AGGGAAAACAGAGAAACTCTGTTGTGTTACCTGGAAACCTGGGGCCCTCTCATC 757	Db	301 AGGGAAAACAGAGAAACTCTGTTGTGTTACCTGGAAACCTGGGGCCCTCTCATC 360							
Qy	758 CGGGAGGCAACAGAGAGGGCTTAACTCTGTTGCTCTGCTGCTGCTGCTGCTG 817	Db	361 CGGGAGGCAACAGAGAGGGCTTAACTCTGTTGCTCTGCTGCTGCTGCTGCTG 420							
Qy	818 TCCCTGGAGGATGAGACATCTAGGATCACTGGCTCTGCTGCTGCTGCTGCTG 877	Db	421 TCCCTGGAGGATGAGACATCTAGGATCACTGGCTCTGCTGCTGCTGCTGCTG 480							
Qy	878 TCAAGCGGCCCTCACCTTCCCTCACTCCAGGCCCTGGTGAACCTTCTGCTGCG 937	Db	481 TCAAGCGGCCCTCACCTTCCCTCACTCCAGGCCCTGGTGAACCTTCTGCTGCG 540							
Qy	938 GATGACATCTCTGCTGCTACTAAGGAGCCCTGTCTGCTGAGGACCACTGAGGT 997	Db	541 GATGACATCTCTGCTGCTACTAAGGAGCCCTGTCTGCTGAGGACCACTGAGGT 600							
Qy	998 GGCAAGGATAACCCCTACCTGTACTGCTGAGGACCACTCACTGGAAAGGTG 1057	Db	601 GGCAAGGATAACCCCTACCTGTACTGCTGAGGACCACTCACTGGAAAGGTG 660							
Qy	1058 GACAGCTCCCTCTGTTCTGAAAGCTGCCAACGGGGAGAGCTCTCTAGTGGGT 1117	Db	661 GACAGCTCCCTCTGTTCTGAAAGCTGCCAACGGGGAGAGCTCTCTAGTGGGT 720							
Qy	1118 CTCCGGAGTCCCTAGCTCTACATCAGCTGATGAACTGGGTCTCTTGGAT 1177	Db	301 AGGGAGAAGCAGGAACTCTGTTACCTGGAAACCTGGGGCCCTCTCATC 360							
DEFINITION	Sequence 4 from Patent WO02055707.	LCUS	AX572845							
DEFINITION	Sequence 1 from Patent WO02055707.	DEFINITION	Sequence 1 from Patent WO02055707.							
ACCESSION	AX572845	ACCESSION	AX572845							
ACCESSION	AX572845	ACCESSION	AX572845							
VERSION	AX572845.1	VERSION	AX572845.1							
KEYWORDS		KEYWORDS								
SOURCE	Homo sapiens (human)	SOURCE	Homo sapiens (human)							
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE	Holland,S.J., Menderhall,M.K., Pardo,J., Spencer,C., Fu,A.C., Luo,Y., Payan,D.G., Mancebo,H.S., Wu,J., Zhou,X., Shen,M., Liao,X.C. and Sheng,N.	REFERENCE	Holland,S.J., Menderhall,M.K., Pardo,J., Spencer,C., Fu,A.C., Luo,Y., Payan,D.G., Mancebo,H.S., Wu,J., Zhou,X., Shen,M., Liao,X.C. and Sheng,N.							
AUTHORS		AUTHORS								
TITLE		TITLE								
JOURNAL		JOURNAL								
FEATURES		FEATURES								
ORIGIN		ORIGIN								
Qy	1 .786 /note="unnamed protein product"	Qy	1 .786 /note="unnamed protein product"							
Db	/codon_start=1	Db	/protein_id="CAD58542.1"							
Db	/db_xref="GI:26004936"	Db	/db_xref="REMBL:CAD58542"							
Db	/translation="MGSPLSPRSLSPLSSLSVQGQPVTMEAERSKATAVALGSFPPRIGPAELSIRGLPRLTIVSEGDWWTIUSEVSREYNIPSPVHAKVSHGWILBGLSREKAEPULLPCNPFGAFLRESPQRGSLSPSLSPRSRASWSDRERHYRLCHNLGWLISPRITFPSPQALVHDYHSEADDCCLKEPCVLRQAGLPKGDIPLPVTVQRIPLNWKEIDLSLLPSSEAATGEESLSEGLESFSYIISNDEAVSLDDA"	Db	/mo_type="unassigned DNA"							
Db	/db_xref="taxon:9606"	Db	/db_xref="taxon:9606"							
Db	1 .786	Db	1 .786							
Db	/note="unnamed protein product"	Db	/note="unnamed protein product"							
Db	/protein_id="CAD58542.1"	Db	/protein_id="CAD58542.1"							
Db	/db_xref="REMBL:CAD58542"	Db	/db_xref="REMBL:CAD58542"							
Db	/translation="MGSPLSPRSLSPLSSLSVQGQPVTMEAERSKATAVALGSFPPRIGPAELSIRGLPRLTIVSEGDWWTIUSEVSREYNIPSPVHAKVSHGWILBGLSREKAEPULLPCNPFGAFLRESPQRGSLSPSLSPRSRASWSDRERHYRLCHNLGWLISPRITFPSPQALVHDYHSEADDCCLKEPCVLRQAGLPKGDIPLPVTVQRIPLNWKEIDLSLLPSSEAATGEESLSEGLESFSYIISNDEAVSLDDA"	Db	/translation="MGSPLSPRSLSPLSSLSVQGQPVTMEAERSKATAVALGSFPPRIGPAELSIRGLPRLTIVSEGDWWTIUSEVSREYNIPSPVHAKVSHGWILBGLSREKAEPULLPCNPFGAFLRESPQRGSLSPSLSPRSRASWSDRERHYRLCHNLGWLISPRITFPSPQALVHDYHSEADDCCLKEPCVLRQAGLPKGDIPLPVTVQRIPLNWKEIDLSLLPSSEAATGEESLSEGLESFSYIISNDEAVSLDDA"							
Db	1 ATGGGAAGTTGCCCTAGGAGAAATCTGCAAGGCCAACGCTTCTGTC 457	Db	1 ATGGGAAGTTGCCCTAGGAGAAATCTGCAAGGCCAACGCTTCTGTC 60							
Db	2 CAAGGCAGGGACCTGTGAACCATGGAAGGAGAACGGCAAGGCCGGCTG 517	Db	2 CAAGGCAGGGACCTGTGAACCATGGAAGGAGAACGGCAAGGCCGGCTG 120							
Db	3 ATGGCAAGTTCCGGAGGTGCCCTAGGAGAACTGGGGAGCCATTGACC 577	Db	3 ATGGCAAGTTCCGGAGGTGCCCTAGGAGAACTGGGGAGCCATTGACC 180							
Db	4 ATGGCAAGTTCCGGAGGTGCCCTAGGAGAACTGGGGAGCCATTGACC 637	Db	4 ATGGCAAGTTCCGGAGGTGCCCTAGGAGAACTGGGGAGCCATTGACC 240							
Db	5 AACATCCCCGGGCAACGTGGCAAAGGCTCCATGGGTGGCTGTATGAGGCCCTGAGC 697	Db	5 AACATCCCCGGGCAACGTGGCAAAGGCTCCATGGGTGGCTGTATGAGGCCCTGAGC 300							
Db	6 AGGGAAAACAGAGAAACTCTGTTGTGTTACCTGGAAACCTGGGGCCCTCTCATC 757	Db	6 AGGGAAAACAGAGAAACTCTGTTGTGTTACCTGGAAACCTGGGGCCCTCTCATC 360							
Db	7 CGGGAGGCAACAGAGAGGGCTTAACTCTGTTGCTCTGCTGCTGCTG 817	Db	7 CGGGAGGCAACAGAGAGGGCTTAACTCTGTTGCTCTGCTGCTGCTG 420							
Db	8 TCAAGCGGCCCTCACCTTCCCTCACTCCAGGCCCTGGTGAACCTTCTGCTGCG 937	Db	8 TCAAGCGGCCCTCACCTTCCCTCACTCCAGGCCCTGGTGAACCTTCTGCTGCG 540							
Db	9 TCCCTGGAGGATGAGACATCTAGGATCACTGGCTCTGCTGCTGCTG 997	Db	9 TCCCTGGAGGATGAGACATCTAGGATCACTGGCTCTGCTGCTGCTG 480							
Db	10 TCAAGCGGCCCTCACCTTCCCTCACTCCAGGCCCTGGTGAACCTTCTGCTGCG 1057	Db	10 TCAAGCGGCCCTCACCTTCCCTCACTCCAGGCCCTGGTGAACCTTCTGCTGCG 600							
Db	11 AACATCCCCGGGCAACGTGGCAAAGGCTCCATGGGTGGCTGTATGAGGCCCTGAGC 1117	Db	11 AACATCCCCGGGCAACGTGGCAAAGGCTCCATGGGTGGCTGTATGAGGCCCTGAGC 720							
Db	12 AGGGAGAAGCAGGAAACTCTGTTGTGTTACCTGGAAACCTGGGGCCCTCTCATC 1177	Db	12 AGGGAGAAGCAGGAAACTCTGTTGTGTTACCTGGAAACCTGGGGCCCTCTCATC 360							





/product="Src-like adaptor protein-2 splice isoform"  
 /protein\_id="AAI38198\_1"  
 /db\_xref="GI:1751923"  
 /translation="MCSLPSRKSLPSLSSVQGPVMEARSKATVALAGSFP  
 AGGAELSLRGEPLITIVSEGDWNTVLSSESGRENTIPSHVAKSHWYEGISRE  
 KAEELLILPQNGCAGLIREQCTRSYSSEVRLSPASMRTRHRYTRHCLDNGLYI  
 SRLTPPSLQLVLDHYSEGWAAPWQGTTPTCDCAEDTQLERAGQLLPV"

ORIGIN

Query	Match	Score	Length
	Best Local Similarity	93.3%	737;
	Pred. No.	4.7e-16	
	0; Mismatches	2;	
	Indels	51;	Gaps 2;
	Matches	734;	Conservative
Qy	398 ATGGGAAGTGCCAGCGAAGAAAATCTGCCAACGCTTGAAGTTCTCTCTC	5.9%	Score 661.8; DB 9;
Db	1 ATGGGAAGTGCCAGCGAAGAAAATCTGCCAACGCTTGAAGTTCTCTCTC	4.57	Length 737;
Qy	458 CAAGGCCAGGACCTGTGACATGCCAGGAGAACATTACTCTGAGCTGGCG	5.17	
Db	61 CAAGGCCAGGACCTGTGACATGCCAGGAGAACATTACTCTGAGCTGGCG	5.17	
Qy	518 GGCAGTTCCGGCAGGGGCCGGCCAGGCGCCAGGAGCTGGCTGAGACTGACC	5.77	
Db	121 GGAGATTCGGCAAGGAGCTGGCTGAGCTGGCTGAGCTGGCTGAGCTGACC	5.77	
Qy	578 ATCGCTCTGAGGATGGAACTGTGAGGAGCTGGCTGAGCTGGCTGAGGAGAT	6.37	
Db	181 ATCGCTCTGAGGATGGAACTGTGAGGAGCTGGCTGAGCTGGCTGAGGAGAT	6.37	
Qy	638 AACATCCCAGCGTCAACGGGGAAAGTCTCCCATGGTGGCTGATGAGGGCTGAGC	6.97	
Db	241 AACATCCCAGCGTCAACGGGGAAAGTCTCCCATGGTGGCTGATGAGGGCTGAGC	6.97	
Qy	698 AGGGAAAGAGGAGAACTGCTGGTACCTGGAAACCTGGAACTGGGACCTCTCTC	7.57	
Db	301 AGGGAAAGAGGAGAACTGCTGGTACCTGGAAACCTGGAACTGGGACCTCTCTC	7.57	
Qy	758 CGGGAGAGCCAGGACCCGAGGAGGCTGCTGAGCTGGCTGAGCTGGCTGAGCTGG	8.17	
Db	361 CGGGAGAGCCAGGACCCGAGGAGGCTGCTGAGCTGGCTGAGCTGGCTGAGCTGG	8.17	
Qy	818 TCTCTGGGACCGGATCAAGGATCACTGGCTCCCTGACAAATGGCTGGCTGACATC	8.77	
Db	421 TCTCTGGGACCGGATCACTGGCTCCCTGACAAATGGCTGGCTGACATC	8.77	
Qy	878 TCACCCGGCCCTCACCTCCCTCACCTCACCGCCCTGACATCTGAGCTGGCG	9.37	
Db	481 TCACCCGGCCCTCACCTCCCTCACCTCACCGCCCTGACATCTGAGCTGGCG	9.37	
Qy	938 GATGACATCTGCTGCTACTCAAGGGCCCTGTGCTGAGGGCTGCCCCCTCC	9.97	
Db	532 -----	-----	
Qy	998 GCGAAGGATAAACCCCTACCTGTGACTGAGGACACACCTCGAAAGAGCTG	10.57	
Db	551 GCGAAGGATAAACCCCTACCTGTGACTGAGGACACACCTCGAAAGAGCTG	10.57	
Qy	1058 GACAGCTCCCTCTGTTTGAACTGTGAGCTGCCACAGGGAGCTCTCTGAGCTGG	11.17	
Db	611 GACAGCTCCCTCTGTTTGAACTGTGAGCTGCCACAGGGAGCTCTCTGAGCTGG	11.17	
Qy	1118 CTCCGGAGATCCCTCACTGCTTACATCAGCTGATGAA-CGAGGCTGCTCTTGGATGA	11.76	
Db	671 CTCCGGAGATCCCTCACTGCTTACATCAGCTGATGAA-CGAGGCTGCTCTTGGATGA	11.76	
Qy	1177 TGCCCTAG 11.83		
Db	731 TGCCCTAG 7.37		

RESULT 12

AF230986 AF230986 737 bp mRNA linear PRI 21-JAN-2003

DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA, complete cds; alternatively spliced.

ACCESSION AF230986

VERSION AF230986.1 GI:17351922

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

MATERIALS Cloning and characterization of human Src-like adaptor protein 2 and a novel splice isoform, SLAP-2-v

JOURNAL Oncogene 22 (2), 266-273 (2003)

MEDLINE 2415750

PUBMED 12527895

REFERENCE 1 (bases 1 to 737)

AUTHORS Lorito, M.P. and McGlaude, C.J.

TITLE Cloning and characterization of human Src-like adaptor protein 2 and a novel splice isoform, SLAP-2-v

JOURNAL Location/Qualifiers

FEATURES source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="0"  
 /tissue\_type="thymus"  
 1..633  
 /note="SLAP-2-v"  
 /codon\_start=1

CDS

DEFINITION	Mus musculus Src-like-adaptor 2, mRNA (cDNA clone MGC:60811 IMAGE:3000401), complete cds.
ACCESSION	BC052655
VERSION	BC052655.1
KEYWORDS	MGC.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Sciuromorpha; Muridae; Murinae; 1 (bases 1 to 138)
AUTHORS	Straussberg, R.L., Feingold, E.A., Grouse, J.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shanmen, C.M., Schulze, A., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N., Hopkins, R.F., Jordan, H., Moore, T., Max, S.J., Wang, J., Hsieh, P., Diatchenko, L., Matsunaga, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Carrascosa, T.E., Brownstein, M.J., Uzdin, T.B., Toshiyuki, S., Carrinci, P., Prange, C., Raha, S.S., Locutio, L.N.A., Peters, G., Abramson, R.D., Mulhall, S.J., Boosk, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Munro, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahsy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, B.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, Butterfield, D.S., Krzywinski, M.I., Stalski, U., Smialowicz, D.E., Scherzer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE	22388257
PUBMED	12477932
REFERENCE	(bases 1 to 138)
AUTHORS	Straussberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (16-MAY-2003) National Institutes of Health, Mammal Gene Collection (MGC). Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2200 USA
REMARK	NIH/MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Web site: <a href="http://www.nsc.nih.gov/">http://www.nsc.nih.gov/</a>
AUTHORS	Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Jonathan Keller (NCI, USA);
TITLE	Tissue Procurement: Drs. Dennis Taub, Dan Longo (NIA, USA);
JOURNAL	CDNA Library Preparation: Yulan Piao and Minoru Ko (National Institute on Aging, NIH: <a href="http://lgsun.grc.nia.nih.gov/cDNA/">http://lgsun.grc.nia.nih.gov/cDNA/</a> )
REMARK	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
COMMENT	Contact: nisc_mgc@nigri.nih.gov
AUTHORS	Blakesley, R.W., Bouffard, G.G., Brinkley, K., Brinkley, R., Granite, S., Guan, X., Gupta, J., Haghichi, P., Dietrich, N.L., Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Laric, P., Legaspi, Madiro, Q., Masiello, C., Maskeri, B., Mastrian, D., McCloskey, J., McDowell, J., Pearson, R., Stanton, P.J., Thomas, P.J., Touchman, J., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, J., Young, A., Zhang, L.-H., and Green, P.D.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL: <a href="http://image.llnl.gov/">http://image.llnl.gov/</a>
source	/organism="Mus musculus" /mol_type="mRNA" /strain="WT/WT"



FEATURES            Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada  
 source            Location\_Qualifiers  
 1..1348            /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_Xref="Taxon:10090"  
 /dev\_stage="day 15 embryo"  
 282 ..1061  
 /note="SLAP-2"  
 /codon\_start=1  
 /product="Src-like adaptor protein-2"  
 /protein\_id="P038196.1",  
 /db\_Xref="NCI:1735193.9"  
 /translation="MGSLSRGRKNSQSPSPSSGGDQEFSMQRHKVTAVALGSEPA  
 GEQLLPLPNPGAAHLIRESDOTRGCVSYLSVRLSPASWDRHRYQLRDLNGWLYIS  
 PRLLTPSLHALVEHTSELLADIGCPLREPCVOLGPUDKTPPPVTPTSSLNWK  
 LDRSLFLLEPASGEASLLESERLSSYISLALDPLDDA"  
 misc\_feature  
 381 ..552            /note="Region: SH3 domain"  
 552 ..824            /note="Region: SH2 domain"  
 misc\_feature  
 1301 ..1306            /polyA\_signal

ORIGIN

Query Match        47.5%;    Score 561.4;    DB 10;    Length 1348;  
 Best Local Similarity    78.2%;    Pred. No. 3.6e-139;  
 Matches            714;    Conservative 0;    Mismatches 191;    Indels 8;    Gaps 3;

Qy	274	AGGGCCCCAAAGCCCTAACCTGTCCAGGCCAGGCTTGCTCTCGAGAGCTGTCCTCC	333
Db	160	AGGACATGCAAAAGCCCTGACCTGTCCGGTCTGGTCTGATCTCAT	217
Qy	334	CAGCCCTTTGAGACACCAATTCCCTCATGATGTGCTCTAGTGTCTGAGG	393
Db	218	CAAACGCTGTGATGCCAAACCCCTTCCCTTCCAGGTTCAAGCTCAGTGCTGCTG	277
Qy	394	AACAAATGGAAACTGCCAGAGAAGAAATCTCTGCCAGGCCAAAGCTTGAATTCCTC	453
Db	278	ATGTATGGAAATTGTCAGAGGAAACCT--CGAACCCCGGCCAGCTCTCCTC	334
Qy	454	TGTCCTAACGGCCAGGACCTGTGACCATGGAGAAGGAGAACGGCTGGC	513
Db	335	TGTCCTAACGGCCAGGAAACCCTGTCCATGACGCCAGAAAGCACAGTGTCGGC	394
Qy	514	CTCTGGGAGTTCCCGCAGGTTGCCAGGCCAGGCTGTCGACTGGGGCCATT	573
Db	395	CCGGGGAGTTCCCAAGGGAACTGGCAAGTATCTGAGACTCTGGGGCCGT	454
Qy	574	GACCATGCTCTGAGATGGAGACTGTGACCTGCTGTTGAGTCTAGGGAGA	633
Db	455	GACCATCATCTGAGATGGAGACTGTGACCTGCTGTTGAGTCTAGGGAGA	514
Qy	634	GTATAACATCCCAGGTCCAGCTGGCAAGTCTCTCCATGGTGGCTGTAGGGCCT	693
Db	515	GTACACATGGCCAGTGTTGCTAAAGTGCCTCCAGGTGGCTGTAGGGCCT	574
Qy	694	GAGGAGGGAGAACAGGAAACTGTGTTTACCTGGAAACCTGGAGGCTTCT	753
Db	575	GAGGCGGGAGAACGGAAACTGTGTTTACCTGGAAACCTGGAGGCTTCT	634
Qy	754	CATCCGGGAGAACGGCACTACGGATCACTCCCTTACCTGGAAACCCGGAGGCT	813
Db	635	CATCCGGGAGAACGGCACTACGGATCACTCCCTTACCTGGAAACCCGGAGGCT	694
Qy	814	TGCACTCTGGACGGATCAGACACTACGGATCACTCCCTTACCTGGAAACCCGGAGGCT	873
Db	695	TGCACTCTGGACGGATCAGACACTACGGATCACTCCCTTACCTGGAAACCCGGAGGCT	754
Qy	874	CATCTAACCCGCTCACCTCCCTAACCTCACGCCCTGGGACCAATTCTGAGT	933
Db	755	CATCTAACCTCCCTAACCTCACGCCCTGGGACCAATTCTGAGT	814

Search completed: February 19, 2004, 23:23:13  
 Job time : 5187.29 secs

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 19, 2004, 21:15:48 / Search time 1096 Seconds  
(without alignments)  
4585.415 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agcttaggcctcaaggacc...tctttttggatgcccgt 1183

Scoring table: IDENTITY\_NUC

Gapp=10.0 , Gapext=1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04\_\*

1: Geneseqn180s:\*

2: Geneseqn1990s:\*

3: Geneseqm2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002s:\*

7: Geneseqn2003as:\*

8: Geneseqn2003bs:\*

9: Geneseqm2003cs:\*

10: geneseqn2004as:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1183	100.0	1183	6	ABK61465		Abk61465 Human cDN
2	784.4	99.9	2567	6	AAD41989		Aad43980 Human SRC
3	784.4	66.3	786	6	AAI44089		Aai44089 Human mod
4	784.4	66.3	786	6	ABQ74493		Abq74343 Human SRC
5	758.2	64.1	763	6	ABQ98670		Abq98670 Human ORF
6	740.2	62.6	837	3	AAC77202		Aac77202 Human ORF
7	724.2	61.2	1413	6	ABQ99374		Abq99374 Human cod
8	651.8	55.9	661	6	AAI44090		Aai44090 Mouse MAR
9	651.8	55.9	2049	5	AAI74750		Aai74750 DNA encoded
10	561.4	47.5	1348	6	AAI44087		Aai44087 Mouse mod
c 11	405	34.2	603	5	AAI74748		Aai74748 DNA encoded
c 12	395	33.4	445	5	AAI74747		Aai74747 DNA encoded
c 13	348	29.4	444	6	ABQ98659		Abq98659 Human ORF
c 14	341	28.8	875	6	ABQ99151		Abq99151 Human ORF
c 15	191.2	16.2	211	5	AAI870481		Aai870481 DNA encoded
c 16	158.6	13.4	2109	4	AAI802049		Aai802049 DNA encoded
c 17	157.4	13.3	2665	6	ABL65189		Abi65189 Lung cancer
c 18	157.4	13.3	2665	6	ABK83738		Abk83738 Human cDN
c 19	157.4	13.3	2665	7	ACB81091		Acb81091 Human SRC
c 20	157.4	13.3	3756	7	ABX62975		Abx62975 Human actin
c 21	141.8	12.0	432	4	AAI12879		Aai12879 Probe #28
c 22	141.8	12.0	432	4	ABM4580		Abm4580 Human fibroblast
c 23	141.8	12.0	432	4	AAI34236		Aai34236 Probe #29

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1183	100.0	1183	6	ABK61465		Abk61465 Human cDN
2	784.4	99.9	2567	6	AAD41989		Aad43980 Human SRC
3	784.4	66.3	786	6	AAI44089		Aai44089 Human mod
4	784.4	66.3	786	6	ABQ74493		Abq74343 Human SRC
5	758.2	64.1	763	6	ABQ98670		Abq98670 Human ORF
6	740.2	62.6	837	3	AAC77202		Aac77202 Human ORF
7	724.2	61.2	1413	6	ABQ99374		Abq99374 Human cod
8	651.8	55.9	661	6	AAI44090		Aai44090 Mouse MAR
9	651.8	55.9	2049	5	AAI74750		Aai74750 DNA encoded
10	561.4	47.5	1348	6	AAI44087		Aai44087 Mouse mod
c 11	405	34.2	603	5	AAI74748		Aai74748 DNA encoded
c 12	395	33.4	445	5	AAI74747		Aai74747 DNA encoded
c 13	348	29.4	444	6	ABQ98659		Abq98659 Human ORF
c 14	341	28.8	875	6	ABQ99151		Abq99151 Human ORF
c 15	191.2	16.2	211	5	AAI870481		Aai870481 DNA encoded
c 16	158.6	13.4	2109	4	AAI802049		Aai802049 DNA encoded
c 17	157.4	13.3	2665	6	ABL65189		Abi65189 Lung cancer
c 18	157.4	13.3	2665	6	ABK83738		Abk83738 Human cDN
c 19	157.4	13.3	2665	7	ACB81091		Acb81091 Human SRC
c 20	157.4	13.3	3756	7	ABX62975		Abx62975 Human actin
c 21	141.8	12.0	432	4	AAI12879		Aai12879 Probe #28
c 22	141.8	12.0	432	4	ABM4580		Abm4580 Human fibroblast
c 23	141.8	12.0	432	4	AAI34236		Aai34236 Probe #29

## ALIGNMENTS

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1183	100.0	1183	6	ABK61465		Abk61465 standard; cDNA; 1183 BP.
2	784.4	99.9	2567	6	AAI44089		Aai44089 Human cDN
3	784.4	66.3	786	6	AAI44089		Aai44089 Human mod
4	784.4	66.3	786	6	ABQ74493		Abq74343 Human SRC
5	758.2	64.1	763	6	ABQ98670		Abq98670 Human ORF
6	740.2	62.6	837	3	AAC77202		Aac77202 Human ORF
7	724.2	61.2	1413	6	ABQ99374		Abq99374 Human cod
8	651.8	55.9	661	6	AAI44090		Aai44090 Mouse MAR
9	651.8	55.9	2049	5	AAI74750		Aai74750 DNA encoded
10	561.4	47.5	1348	6	AAI44087		Aai44087 Mouse mod
c 11	405	34.2	603	5	AAI74748		Aai74748 DNA encoded
c 12	395	33.4	445	5	AAI74747		Aai74747 DNA encoded
c 13	348	29.4	444	6	ABQ98659		Abq98659 Human ORF
c 14	341	28.8	875	6	ABQ99151		Abq99151 Human ORF
c 15	191.2	16.2	211	5	AAI870481		Aai870481 DNA encoded
c 16	158.6	13.4	2109	4	AAI802049		Aai802049 DNA encoded
c 17	157.4	13.3	2665	6	ABL65189		Abi65189 Lung cancer
c 18	157.4	13.3	2665	6	ABK83738		Abk83738 Human cDN
c 19	157.4	13.3	2665	7	ACB81091		Acb81091 Human SRC
c 20	157.4	13.3	3756	7	ABX62975		Abx62975 Human actin
c 21	141.8	12.0	432	4	AAI12879		Aai12879 Probe #28
c 22	141.8	12.0	432	4	ABM4580		Abm4580 Human fibroblast
c 23	141.8	12.0	432	4	AAI34236		Aai34236 Probe #29

RESULT 1

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1183	100.0	1183	6	ABK61465		Abk61465 standard; cDNA; 1183 BP.
2	784.4	99.9	2567	6	AAD41989		Aad43980 Human SRC
3	784.4	66.3	786	6	AAI44089		Aai44089 Human mod
4	784.4	66.3	786	6	ABQ74493		Abq74343 Human SRC
5	758.2	64.1	763	6	ABQ98670		Abq98670 Human ORF
6	740.2	62.6	837	3	AAC77202		Aac77202 Human ORF
7	724.2	61.2	1413	6	ABQ99374		Abq99374 Human cod
8	651.8	55.9	661	6	AAI44090		Aai44090 Mouse MAR
9	651.8	55.9	2049	5	AAI74750		Aai74750 DNA encoded
10	561.4	47.5	1348	6	AAI44087		Aai44087 Mouse mod
c 11	405	34.2	603	5	AAI74748		Aai74748 DNA encoded
c 12	395	33.4	445	5	AAI74747		Aai74747 DNA encoded
c 13	348	29.4	444	6	ABQ98659		Abq98659 Human ORF
c 14	341	28.8	875	6	ABQ99151		Abq99151 Human ORF
c 15	191.2	16.2	211	5	AAI870481		Aai870481 DNA encoded
c 16	158.6	13.4	2109	4	AAI802049		Aai802049 DNA encoded
c 17	157.4	13.3	2665	6	ABL65189		Abi65189 Lung cancer
c 18	157.4	13.3	2665	6	ABK83738		Abk83738 Human cDN
c 19	157.4	13.3	2665	7	ACB81091		Acb81091 Human SRC
c 20	157.4	13.3	3756	7	ABX62975		Abx62975 Human actin
c 21	141.8	12.0	432	4	AAI12879		Aai12879 Probe #28
c 22	141.8	12.0	432	4	ABM4580		Abm4580 Human fibroblast
c 23	141.8	12.0	432	4	AAI34236		Aai34236 Probe #29

RESULT 2

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1183	100.0	1183	6	ABK61465		Abk61465 standard; cDNA; 1183 BP.
2	784.4	99.9	2567	6	AAD41989		Aad43980 Human SRC
3	784.4	66.3	786	6	AAI44089		Aai44089 Human mod
4	784.4	66.3	786	6	ABQ74493		Abq74343 Human SRC
5	758.2	64.1	763	6	ABQ98670		Abq98670 Human ORF
6	740.2	62.6	837	3	AAC77202		Aac77202 Human ORF
7	724.2	61.2	1413	6	ABQ99374		Abq99374 Human cod
8	651.8	55.9	661	6	AAI44090		Aai44090 Mouse MAR
9	651.8	55.9	2049	5	AAI74750		Aai74750 DNA encoded
10	561.4	47.5	1348	6	AAI44087		Aai44087 Mouse mod
c 11	405	34.2	603	5	AAI74748		Aai74748 DNA encoded
c 12	395	33.4	445	5	AAI74747		Aai74747 DNA encoded
c 13	348	29.4	444	6	ABQ98659		Abq98659 Human ORF
c 14	341	28.8	875	6	ABQ99151		Abq99151 Human ORF
c 15	191.2	16.2	211	5	AAI870481		Aai870481 DNA encoded
c 16	158.6	13.4	2109	4	AAI802049		Aai802049 DNA encoded
c 17	157.4	13.3	2665	6	ABL65189		Abi65189 Lung cancer
c 18	157.4	13.3	2665	6	ABK83738		Abk83738 Human cDN
c 19	157.4	13.3	2665	7	ACB81091		Acb81091 Human SRC
c 20	157.4	13.3	3756	7	ABX62975		Abx62975 Human actin
c 21	141.8	12.0	432	4	AAI12879		Aai12879 Probe #28
c 22	141.8	12.0	432	4	ABM4580		Abm4580 Human fibroblast
c 23	141.8	12.0	432	4	AAI34236		Aai34236 Probe #29

RESULT 3

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1183	100.0	1183	6	ABK61465		Abk61465 standard; cDNA; 1183 BP.
2	784.4	99.9	2567	6	AAD41989		Aad43980 Human SRC
3	784.4	66.3	786	6	AAI44089		Aai44089 Human mod
4	784.4	66.3	786	6	ABQ74493		Abq74343 Human SRC
5	758.2	64.1	763	6	ABQ98670		Abq98670 Human ORF
6	740.2	62.6	837	3	AAC77202		Aac77202 Human ORF
7	724.2	61.2	1413	6	ABQ99374		Abq99374 Human cod
8	651.8	55.9	661	6	AAI44090		Aai44090 Mouse MAR
9	651.8	55.9	2049	5</td			

NOVX, a NOVX variant (differing by no more than 15%), the nucleotide encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14, NOV15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it, and antibody against it, are useful for treating or preventing (e.g., by gene therapy) a NOVX-associated disorder in humans, e.g., cardiomypathy, thetherosclerosis, a disorder related to cell signal processing and metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide and nucleic acids are also useful for determining the presence of predisposition to the diseases. The NOVX nucleic acid and polypeptide are especially useful in therapeutic or prophylactic applications for disorders associated with aberrant NOVX expression or activity, e.g., cancers (e.g., adenocarcinoma, lymphoma, prostate cancer or uterus cancer), immune response, graft-versus-host disease, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension, congenital heart defects, multiple sclerosis, inflammation or Albright hereditary osteodystrophy and many other diseases listed in the specification. The DNA encoding the protein is useful in gene therapy for treating the conditions. This is also useful in detection assays, chromosome mapping, tissue typing, diagnostic or prognostic assays, or for developing a powerful assay system for functional analysis of various human disorders, as well as in diagnostic applications. The present sequence encodes a NOVX protein.

XX Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

CC	Qy	661	CAAAGTCTCCCATTGGGGCTGTATGAGGGAGAAAGCAGGGAAACTGCT	720
CC	Db	661	CAAAGTCTCCCATTGGGGCTGTATGAGGGAGAAAGCAGGGAAACTGCT	720
CC	Qy	721	GTTGTTACCTGGAAACCTGAGGGCCCTCTCATCGGAGAGCCAGGGAGG	780
CC	Db	721	GTTGTTACCTGGAAACCTGAGGGCCCTCTCATCGGAGAGCCAGGGAGG	780
CC	Qy	781	CTCTTACTCTCTGTCAGTCAGGCCCTGCTCATCTGGGACGGATCAGCA	840
CC	Db	781	CTCTTACTCTCTGTCAGTCAGGCCCTGCTCATCTGGGACGGATCAGCA	840
CC	Qy	841	CAGGATCACAGCCTGAATGGCTGCTGATCTCACCGGCCTCACCTGCT	900
CC	Db	841	CAGGATCACAGCCTGAATGGCTGCTGATCTCACCGGCCTCACCTGCT	900
CC	Qy	901	ACTCCAGGCCCTGGACCAATTACTGACTGGGATGACATCTGCTGCT	960
CC	Db	901	ACTCCAGGCCCTGGACCAATTACTGACTGGGATGACATCTGCTGCT	960
CC	Qy	961	GGAGGCCCTGTCCTGCTGAGGGCTGGCCCTCTGGAGGATATACCC	1020
CC	Db	961	GGAGGCCCTGTCCTGCTGAGGGCTGGCCCTCTGGAGGATATACCC	1020
SQ	Qy	1021	GACTGTGAGGAGAACCTCAACTGTAAGAGCTGAGAGCTCCCTGTTCTGA	1080
SQ	Db	1021	GACTGTGAGGAGAACCTCAACTGTAAGAGCTGAGAGCTCCCTGTTCTGA	1080
Matches 1183 ; Best Local Similarity 100.0% ;保守性 0; Mismatches 0; Indels 0; Gaps 0;	Qy	1081	AGTGCTCACAGGGAGAGCTCTCTGAGGCTCTGGGACTCCCTCAGCTTA	1140
Db	1081	AGTGCTCACAGGGAGAGCTCTCTGAGGCTCTGGGACTCCCTCAGCTTA	1140	
Qy	1141	CATCAGGCTGAATGAGGGCTGTCMCTTGGATGATGCCPAG	1183	
Db	1141	CATCAGGCTGAATGAGGGCTGTCMCTTGGATGATGCCPAG	1183	
RESULT 2				
		AAD43980	standard	cDNA; 2567 BP.
	ID	AAD43980		
	XX	AAD43980:		
	AC	AAD43980:		
	XX		DT 13 - DEC - 2002	(first entry)
	XX		Human Src-Like Adapter Protein-2 (hSLAP-2)	CDNA.
	DE			
	XX		Human; SH2 / SH3-domain-containing adapter; Src-Like Adapter Protein-2;	
	KW		SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;	
	KW		neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;	
	KW		psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;	
	KW		Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;	
	KW		multiple sclerosis; asthma; acute respiratory distress syndrome;	
	KW		pulmonary disorder; dermatological; neuroprotective; gene; ss.	
	XX		Homo sapiens.	
	XX		Key	Location/Qualifiers
	FT		CDS	415..1200
	FT		/tag=	a "Human SLAP-2"
	XX		/product=	"Human SLAP-2"
	XX		WO200242457-A1.	
	PD		30-MAY-2002.	
	XX		20-NOV-2001;	2001WO-US043367.
	XX		22-NOV-2000;	2000US-0252545P.
	PR			
	PA		(BRIM ) BRISTOL-MYERS SQUIBB CO.	
	XX		Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ,	Kanner SB;
	P1			





Query	Subject	Match	Score	Length	DB	6;	DB	786;	
		Best Local Similarity	Pred. No.	No.	1-9e-202;		Indels	Gaps	0
		Matches	Conservative	0;	Mismatches	1;			
Qy	Db	398	ATGCCAAGTGTCCAGCAAGAAATTCCTGCAACGCCAACAGCTTGAGTTCTCTGTC	457					
Qy	Db	458	CAAGCCAGGGACCTGTGACCATGAAAGCAGAGAAGCAAGCCACAGCCGCTGCC	517					
Qy	Db	61	CAAGGCCAGGGACCTGTGACCATGAAAGCAGAGAAGCAAGCCACAGCCGCTGCC	120					
Qy	Db	518	GGCAATTCCGGAGGTGGCCGGCGGAGGCTTGCTGAGCTGGGGAGGCAATGACC	577					
Qy	Db	121	GGCAGTTCCGGAGGTGGCCGGCGGAGGCTTGCTGAGCTGGGGAGGCAATGACC	180					
Qy	Db	578	ATCGCTCTGAGGATGGAGACTGGTGGAGGACTGTGGAGGGTGTGGTGAAGCTAT	637					
Qy	Db	181	ATCGCTCTGAGGATGGAGACTGGTGGAGGACTGTGGAGGGTGTGGTGAAGCTAT	240					
Qy	Db	638	AACATCCCCAGGTCACCTGGGAAAGTCCTCCATGGTTGTATGGGGTGTAGCTGG	697					
Qy	Db	241	AACATCCCCAGGTCACCTGGGAAAGTCCTCCATGGTTGTATGGGGTGTAGCTGG	300					
Qy	Db	698	AGGGAGAAGCCAGGAACTGCTTGTACTCTGGAAACCCCTGGAGGGCCCTCTCTCATC	757					
Qy	Db	301	AGGGAGAAGCCAGGAAACTGCTTGTACTCTGGAAACCCCTGGAGGGCCCTCTCTCATC	360					
Qy	Db	758	CGGGAGGAGCAAGGAGGGCTTACTCTGTGAGCTGGCCCTGAGCTGGGGCTGAGCT	817					
Qy	Db	361	CGGGAGGAGCAAGGAGGGCTTACTCTGTGAGCTGGGGCTGAGCTGGGGCTGAGCT	420					
Qy	Db	818	TCCCTGGACCGGATCAAGACATCAAGGATCCACTGCTTGCATAATGGTGTAGATC	877					
Qy	Db	421	TCCCTGGACCGGATCAAGACATCAAGGATCCACTGCTTGCATAATGGTGTAGATC	480					
Qy	Db	878	TCACCGGGCTCACTTCCCTCTACTCAGGGCTTGTGAGCTGAGGACATTACTCTGAGG	937					
Qy	Db	481	TCACTGGGCTCACCTTCCCTCTACTCAGGGCTTGTGAGCTGAGGACATTACTCTGAGG	540					
Qy	Db	938	GATGAGATATACCCACTGTGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG	997					
Qy	Db	541	GATGAGATATACCCACTGTGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG	600					
Qy	Db	998	GGCAAGGATAATACCCACTGTGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG	105					
Qy	Db	601	GGCAAGGATAATACCCACTGTGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG	660					
Qy	Db	1058	GACAGTCCTCCCTGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGG	111					
Qy	Db	661	GACAGTCCTCCCTGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGG	720					
Qy	Db	1118	CTCGGGAGCTCCCTGAGCTCCCTGAGCTCCCTGAGCTCCCTGAGCTCCCTGAGCTG	117					

**RESULT 5**

Db	721	CTCGGGAGTCCTCAGTTACATAGCCCTGAAATGAGGCTGTCTTGGAT	780
ABQ98670	1178	GCCTAG	
ID	ABQ98670	standard; DNA; 763 BP.	
XX	AC	ABQ98670;	
XX	DT	04-NOV-2002 (first entry)	
XX	DE	Human ORF477 coding sequence.	
XX	FN	Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;	
XX	FN	Antiinflammatory; Gene therapy; human; ORFX; platelet;	
XX	FN	human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;	
XX	FN	cancer; cardiovascular disease; allergy; autoimmune disease;	
XX	FN	wound healing; blood coagulation disorder; inflammatory disorder; ds.	
XX	OS	Homo sapiens.	
XX	PN	US20000822006-A1.	
XX	PD	27-JUN-2002.	
XX	PP	30-MAY-2001; 2001US-00867550.	
XX	PR	30-MAY-2000; 2000US-0208427P.	
XX	PA	(LEAC/) LEACH M. D.	
PA	PA	(MEHR/ ) MEHRABIAN F.	
PA	PA	(CONL/ ) CONLEY P. B.	
PA	PA	(TOPP/ ) TOPPER J. N.	
PA	PA	(LAWD/ ) LAW D.	
XX	PI	Leach MD, Mehrabian F, Conley PB, Topper JN, Law D;	
XX	DR	WPI; 2002-62655/67.	
DR	P-PSDB; ABP64107.		
XX	PT	New polypeptide designated ORFX are present in human atherogenic cells	
PT	PT	and are useful to prevent and treat ORFX-associated disorders including	
PT	PT	cancer, allergy, wound healing or autoimmune, cardiovascular or	
PT	PT	inflammatory disease.	
XX	BS	Claim 2: SEQ ID NO 953; 78pp; English.	
XX	CC	The present invention relates to novel human ORFX polypeptides and their	
CC	CC	coding sequences (ABP6363-ABP94681 and ABQ98194-ABQ99267). The sequences	
CC	CC	were discovered in human atherogenic cells, in particular in platelets	
CC	CC	and human umbilical vein endothelial cells (HUVEC) and are expressed in	
CC	CC	many other tissues as well. Atherogenic cells are cells which have the	
CC	CC	potential to develop atherosclerotic plaques. The ORFX polypeptides and	
CC	CC	nucleic acids are useful for treating or preventing a pathological	
CC	CC	condition associated with an ORFX-associated disorder, e.g. cancer,	
CC	CC	cardiovascular disease, allergy, autoimmune disease, wound healing, blood	
CC	CC	coagulation disorders or inflammatory disorders. Note: The sequence data	
CC	CC	for this patent did not form part of the printed specification, but was	
CC	CC	obtained in electronic format directly from the USPTO web site at	
XX	SQ	seqdata.uspto.gov/sequence.html?DocID=20020082206	
Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;			
Query Match	Score 758.2;	DB 6;	
Best Local Similarity	Length 763;		
Matches 760;	Pred. No. 2.4e-18;		
Conservative	0; Missmatches 3;		
	Indels 0; Gaps 0;		
113 CTAAAGGCCATGCCGCCAGCTTACATAGCCCTGAAATGAGGCTGTCTTGGAT	172		

1 CTATGGACATGGAGCTGACTCATOCCCTGGTACAAACTGTGACTCGAGAGAT 60  
 Db 173 GCTGAGCTTACCAAAACAAACCTAGGCTCCTGAAGATCTCCAGGTGAGAGT 232  
 Qy 61 GCTGAGCTTACCAAAACAAACCTAGGCTCCTGAAGATCTCCAGGTGAGAGT 120  
 Db 233 TCGGGGTGTCCTAGGACAAAGGACACTGGAGACTTCCAGAAGGGCCCCAAAGCCATAA 292  
 Qy 121 TCGGATOTCPAGGACAAAGGACACTGGCAACTTCAGAGAACCTGGCTTA 180  
 Db 293 CCTGTCAGCAGAGATGGCTCTCACAGAGCTGCTTCCAAAGCCTTATGACAAAC 352  
 Qy 181 CCTGTCAGCAGAGATGGCTCTCACAGAGCTGCTTCCAAAGCCTTATGACAAAC 240  
 Db 353 CAATTTCCTCGATGATGTGATGTGATGTGATGTGATGTGATGTGCC 412  
 Qy 241 CAATTTCCTCGATGATGTGATGTGATGTGATGTGCC 300  
 Db 413 AGCGAAAGAAATCTCTCCAAGCCCAGCTTGAGTTGAGTCCTGTCTGTC 472  
 Qy 301 AGCGAAAGAAATCTCTCCAAGCCCAGCTTGAGTTGAGTCCTGTCTGTC 360  
 Db 473 GTGACCATGGAAAGAGAGAAAGAACGCCAACGCCAACGCCAACGCCAAC 532  
 Db 361 GTGACCATGGAAAGAGAAAGAACGCCAACGCCAACGCCAACGCCAAC 420  
 Qy 533 GGTGGCCGGAGCTGCTGCTGAGACTGGAGCTGGAGCTGGAGCATGGGGAGGCCT 592  
 Db 421 GGTCGCCGGAGCTGCTGAGACTGGAGCATGGGGAGCTGGGGAGCTGGGGAT 480  
 Qy 593 GGAGACTGTTGGAGGCTGCTGTGAGTCTGGCAGAGTATAACATCCAGCGCTC 652  
 Db 481 GGAACTGTGGAGGTGTTGACTCAGGAGACTTACATCCACCGTC 540  
 Qy 653 CACGGGGCAAGTCTCCATGGGTGGCTATGAGGGCCCTGAGGGAGAAAGGAGAG 712  
 Db 541 CACGGGGCAAAAGTCTCCATGGGTGGCTATGAGGGCCCTGAGGGAGAAAGGAG 600  
 Qy 713 GAATGCTGTGTGTACCTGGAAACCCCTGGAGGGCTTCCCTATCCGGAAAGCGACC 772  
 Db 601 GAACTGCTGTGTGTACCTGGAAACCCCTGGAGGGCTTCCCTATCCGGAAAGCGACC 660  
 Qy 773 AGGAGGGCTTFACTCTGTGAGTCGGCTTAGGCCCTGCACTCTGGAGCCGATC 832  
 Db 661 AGGAGGGCTTFACTCTGTGAGTCGGCTTAGGCCCTGCACTCTGGAGCCGATC 720  
 Qy 833 AGACACTAGGATCCACTGCTTGACATGGCTGGCTGACA 875  
 Db 721 AGACACTAGGATCCACTGCTTGACATGGCTGGCTGACA 763

RESULT 6  
 AAC77202 standard; cDNA; 837 BP.  
 ID AAC77202;  
 AC AAC77202;  
 DT 08-FEB-2001 (first entry)  
 XX Human ORF2757 polynucleotide sequence SEQ ID NO:5513.  
 KW Human; open reading frame; ORF; detection; cytosolic; hepatotropic; vulnery; anticonvulsant; thrombolytic; antiparkinsonian; nootropic; neuroprotective; antihypertensive; dermato logical; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antineoplastic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypertension; hypothyroidism; cholesterol ester storage; systemic lupus erythematosus; severe combined immunodeficiency (SCID); AIDS; septicemia; aplastic anemia; burns, wounds, bone and cartilage damage; nocturnal haemoglobinuria; antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive; ss.  
 KW Homo sapiens.  
 XX WO200058473-A2.  
 XX PD 05-OCT-2000.  
 XX PF 2000WO-US00B621.  
 XX PR 31-MAR-1999; 99US-012760P.  
 XX PR 02-APR-1999; 99US-012763P.  
 XX PR 05-APR-1999; 99US-012772P.  
 XX PR 30-MAR-2000; 2000US-00540783.  
 XX (CURA-) CURAGEN CORP.  
 XX PA Shimkets RA, Leach M;  
 XX PI WPI; 2000-602362/57.  
 DR P-2DB; AAB42993.  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 Claim 5; Page 4692-4693; 5507pp; English.  
 XX AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORF6 open reading frames 1 to 3161. The ORF6 sequences have activities such as: cytosolic; hepatotropic; vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticonvulsant; antiarthritic; immunosuppressant; immunostimulant; cardiotonic; thrombolytic; coagulant; vasotropics; antidiabetic; hyporesponsive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an CC associated disorder. The nucleic acids can be used to express ORF6 proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX Sequence 837 BP; 176 A; 245 C; 160 T; 0 U; 2 Other;  
 SQ Query Match 62.6%; Score 740.2; DB 3; Length 837;  
 Best Local Similarity 99.6%; Pred. No. 1.9e-19;  
 Matches 742; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 439 AAGCTGATGTTCCCTGTGTCAGGCCAGGGACCTGTGACCATGGAGAGAAAGAA 498  
 Db 3 AAGCTGATGTTCCCTGTGTCAGGCCAGGGACCTGTGACCATGGAGAGAAAGAA 62  
 Qy 499 GGCCACAGGGCTGGCCCTGGCACTTCCGGAGGTTGCCCTGGCTGCTGCTGAG 558  
 Db 63 GGCCACAGGGCTGGCCCTGGCACTTCCGGAGGTTGCCCTGGCTGCTGCTGAG 122  
 Qy 559 ACTCGGGGCGCCATTGACCATGCTCTGGAGACTGGAGGGCTGCTGCTGAG 618  
 Db 123 ACTCGGGGCGCCATTGACCATGCTCTGGAGACTGGAGGGCTGCTGCTGAG 182  
 Qy 619 AGCTCGAGGGAGAGATAACATGCCAGGGCCAAAGCTGCCATGGCTGGCTGG 678  
 Db 183 AGTCTCAGGGAGAGATAACATGCCAGGGCCAAAGCTGCCATGGCTGGCTGG 242  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;



Db	481	ACCGATCAAGACTACAGGATCCACTCCCTTGACAATTGGCTGTACATCTCACCGC	540	disorders, immunosuppression, myeloproliferative disorders and malfunctions related to the modulation of tyrosine kinases (e.g. breast cancer). The present cDNA sequence encodes a mouse MARS protein.
Qy	885	GCCCTACCTTCCCTCACTCCAGCCCTGGGACCAATTCTCTGCTGGGATGACA	944	CC Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;
Db	541	GCCTOACCTTCCCTCACTCCAGCCCTGGGACCAATTCTCTGCTGGGATGACA	584	Query Match 55.9%; Score 661.8; DB 6; Length 737;
Qy	945	TCTGCTGCCCTACTCAAGGAGCCCTGTGTCAGGGCTCCCTGGGAAGG	1004	Best Local Similarity 93.3%; Pred. No. 3; 3e-169;
Db	585	-----GAGGGCTGCCCTCTGGGAAGG	610	Matches 734; Conservative 0; Mismatches 2; Indels 51; Gaps 2;
Qy	1005	ATATAACCCTAACCTGTAACGTGTGCAAGGACACCACTCAACTCNGAAGG	1064	Qy 3.98 ATGGAAAGTTGCCAGAGAGAAATACTCTGCCAACGCCAGCTTGAGTTCCTCTCTC 457
Db	611	ATATAACCCTAACCTGTAACGTGTGCAAGGACACCACTCAACTCNGAAGG	670	Db 1 ATGGAAAGTTGCCAGAGAGAAATACTCTGCCAACGCCAGCTTGAGTTCCTCTC 60
Qy	1065	CCCTOCTGTTTCTGAACTGCACTCCAGGGAGGAGTCTTCAGTGAGGTCTCGGG	1124	Qy 4.58 CAAGGCCAAGGAGACCTGTGACCATGGAACGAGAACGACAGGCAAGGCCAGAGAGAAGGAACTGGAGCTGGGAGCTGGGAGGAACTGGGAGCATTGCC 517
Db	671	CCCTOCTGTTTCTGAACTGCACTCCAGGGAGGAGTCTTCAGTGAGGTCTCGGG	730	Db 61 CAAGGCCAAGGAGACCTGTGACCATGGAACGAGAACGACAGGCAAGGCCAGAGAGAAGGAACTGGAGCTGGAGCTGGGAGGAACTGGGAGCATTGCC 120
Qy	1125	AGTCCTCACTCTTACATCGCTGAACTGAGGTCTTGGATGATGCTTAG	1183	Qy 5.18 GGCAAGTTCCCGGAGGTGGCCGGAGCTGGCTGAGACTCGGGAGGAACTGGGAGCATTGCC 577
Db	731	AGTCCTCACTCTTACATCGCTGAACTGAGGTCTTGGATGATGCTTAGCTAG	789	Db 121 GGCACTTCCGGAGGTGGCCGGAGCTGGCTGAGACTGGAGCCATTGCC 180
Qy	578	ATCGCTCTCTGAGGATGGAGACTGTGGAGCTGTGGTGAATCTCGGGAGAGATA	637	Qy 5.78 ATCGCTCTCTGAGGATGGAGACTGTGGAGCTGTGGTGAATCTCGGGAGAGATA 637
Db	181	ATCGCTCTCTGAGGATGGAGACTGTGGAGCTGTGGTGAATCTCGGGAGAGATA	240	Db 181 ATCGCTCTCTGAGGATGGAGACTGTGGAGCTGTGGTGAATCTCGGGAGAGATA 240
Qy	638	AACATCCCAAGGTGCACTGGAAAGTCTCCATGGTGTATGAGGCCCTGGC	697	Qy 6.38 AACATCCCAAGGTGCACTGGAAAGTCTCCATGGTGTATGAGGCCCTGGC 697
Db	241	AACATCCCAAGGTGCACTGGAAAGTCTCCATGGTGTATGAGGCCCTGGC	300	Db 241 AACATCCCAAGGTGCACTGGAAAGTCTCCATGGTGTATGAGGCCCTGGC 300
Qy	698	AGGGAAAGCAGAGGAATGGGAACTCCTGGTACCTGGAACTGGGAGGCTCTCATC	757	Qy 6.98 AGGGAAAGCAGAGGAATGGGAACTCCTGGTACCTGGGAGGCTCTCATC 757
Db	301	AGGGAAAGCAGAGGAATCTGGTGTGTTACCTGGAACTGGGAGGCTCTCATC	360	Db 301 AGGGAAAGCAGAGGAATCTGGTGTGTTACCTGGAACTGGGAGGCTCTCATC 360
Qy	758	CGGGAGGAGCCAGACGGAGGGCTACTCTGTAAGTGGCTGACTCTGTCAGTC	817	Qy 7.58 CGGGAGGAGCCAGACGGAGGGCTACTCTGTAAGTGGCTGACTCTGTCAGTC 817
Db	361	CGGGAGGAGCCAGACGGAGGGCTACTCTGTCAGTCAGTCAGTCAGTCAGTC	420	Db 361 CGGGAGGAGCCAGACGGAGGGCTACTCTGTCAGTCAGTCAGTCAGTCAGTC 420
Qy	818	TCCCTGGAGGACGGATCAGACACTAGGATCCACTGCTGAGTGTGACTCTG	877	Qy 8.18 TCCCTGGAGGACGGATCAGACACTAGGATCCACTGCTGAGTGTGACTCTG 877
Db	421	TCCCTGGAGGACGGATCAGACACTAGGATCCACTGCTGAGTGTGACTCTG	480	Db 421 TCCCTGGAGGACGGATCAGACACTAGGATCCACTGCTGAGTGTGACTCTG 480
Qy	878	TCACCGGGCTCACCTCCCTCACTCCAGGGCTCTGGGACCTTACTCTGAGCTGGCG	937	Qy 8.78 TCACCGGGCTCACCTCCCTCACTCCAGGGCTCTGGGACCTTACTCTGAGCTGGCG 937
Db	481	TCACCGGGCTCACCTCCCTCACTCCAGGGCTCTGGGACCTTACTCTGAGCTGGCG	531	Db 481 TCACCGGGCTCACCTCCCTCACTCCAGGGCTCTGGGACCTTACTCTGAGCTGGCG 531
Qy	938	GATGACATCTGGCTGCTCAAGAGGCTCTGGTCCAGAGGCTCTGGCTCTGGCT	997	Qy 9.38 GATGACATCTGGCTGCTCAAGAGGCTCTGGTCCAGAGGCTCTGGCTCTGGCT 997
Db	532	-----GAGGGCTGGCCGCTCTCCCT	550	Db 532 -----GAGGGCTGGCCGCTCTCCCT 550
Qy	998	GGCAAGGAGATAACCCCTACCTGTAAGCTGTGAGGAAACCACTCACTGGAAAGAGGTG	1057	Qy 9.98 GGCAAGGAGATAACCCCTACCTGTAAGCTGTGAGGAAACCACTCACTGGAAAGAGGTG 1057
Db	551	GGCAAGGAGATAACCCCTACCTGTAAGCTGTGAGGAAACCACTCACTGGAAAGAGGTG	610	Db 551 GGCAAGGAGATAACCCCTACCTGTAAGCTGTGAGGAAACCACTCACTGGAAAGAGGTG 610
Qy	1058	GACAGCTCCCTCCCTGTTCTGAGGTGCACTGGGAGGAGTCTCTGAGGTCTCTGAGGT	1117	Qy 10.58 GACAGCTCCCTCCCTGTTCTGAGGTGCACTGGGAGGAGTCTCTGAGGTCTCTGAGGT 1117
Db	611	GACAGCTCCCTCCCTGTTCTGAGGTGCACTGGGAGGAGTCTCTGAGGTCTCTGAGGT	670	Db 611 GACAGCTCCCTCCCTGTTCTGAGGTGCACTGGGAGGAGTCTCTGAGGTCTCTGAGGT 670
Qy	1118	CTCCGGAGAATCCCTCACTACCCCTGATGAATGAGCTCTTGGATGA	1176	Qy 11.18 CTCCGGAGAATCCCTCACTACCCCTGATGAATGAGCTCTTGGATGA 1176
Db	671	CTCCGGAGAATCCCTCACTACCCCTGATGAATGAGCTCTTGGATGA	730	Db 671 CTCCGGAGAATCCCTCACTACCCCTGATGAATGAGCTCTTGGATGA 730
Qy	1177	TGCCPAG 1183		Qy 11.77 TGCCPAG 1183
Db	731	TGCCPAG 737		Db 731 TGCCPAG 737
PS	Claim 9; Page 77; 110PP; English.			
XX	The invention comprises the amino acid and coding sequences of modulator			
CC	of antigen receptor signalling (MARS) proteins. The MARS protein is a			
CC	putative tumour suppressor gene and exhibits structural and sequence			
CC	similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and			
CC	protein sequences of the invention are useful for the treatment of			
CC	myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune			
PS	maligancies			
XX	maligancies			
CC	and myeloproliferative disorders such as myeloid			
PT	malignancies, autoimmune disorders and myeloproliferative disorders.			
XX				
PT	New isolated modulator of antigen receptor signalling protein or its			
PT	fragment, useful for treating malignant disorders such as myeloid			
PT	malignancies, autoimmune disorders and myeloproliferative disorders.			
XX				
PS	Claim 9; Page 77; 110PP; English.			
XX	The invention comprises the amino acid and coding sequences of modulator			
CC	of antigen receptor signalling (MARS) proteins. The MARS protein is a			
CC	putative tumour suppressor gene and exhibits structural and sequence			
CC	similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and			
CC	protein sequences of the invention are useful for the treatment of			
CC	myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune			
PS	maligancies			
XX	maligancies			
CC	and myeloproliferative disorders such as myeloid			
PT	malignancies, autoimmune disorders and myeloproliferative disorders.			
XX				
PT	RESULT 9			
ID	AAS74750			
TD	AAS74750 standard; cDNA; 2049 BP.			
XX				
AC	AAS74750;			
XX				

DT	13-FEB-2002	(first entry)	
XX	DNA encoding novel human diagnostic protein #10554.		
DE			
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.		
KW			
RW			
XX			
OS	Homo sapiens.		
XX			
PN	WO200175067-A2.		
XX			
PD	11-OCT-2001.		
XX			
PR	30-MAR-2001; 2001WO-US008631.		
XX			
PR	31-MAR-2000; 2000US-00540217.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Drmanac RT, Liu C, Tang YR;		
XX			
DR	WPI; 2001-639362/73.		
XX			
P-PSDB	AEG10563.		
XX			
PT	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.		
PT			
PS	Claim 1: SEQ ID NO 10554; 103pp; English.		
XX			
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (II) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) and its binding partners are useful for treating diorrdoses involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/publisted_pct_sequences		
XX	Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;		
Query Match	55.0%; Score 651; DB 5; Length 2049;		
Best Local Similarity	92.5%; Pred. No. 3.66-166;		
Matches	760; Conservative 0; Mismatches 8; Indels 54; Gaps 5;		
XX			
OS	Mouse sp.		
XX			
FH	Location/Qualifiers		
XX			
Key			
CDS			
FT			
FT			
FT			
XX			
PN	WO200242452-A2.		
XX			
PD	30-MAY-2002.		
XX			
PF	26-NOV-2001; 2001WO-CA001662.		
XX			
PR	27-NOV-2000; 2000CA-02324663.		

XX	(HOSP-) HOSPITAL FOR SICK CHILDREN.	Qy	934 GGCGATGAGATCTGCTGCTACTAAGAGGGCTGGCCGT 993
XX	Mcglade JC, Loreto MP;	Db	815 AGCAGATGGCATCGTGTGCCCCCTCAAGGAGCTTGCCACT 874
XX	WPI; 2002-566564/60.	Qy	994 CCCTGGAAAGATAACCCCTAACCTGTAACCTGAACCTGAAAGAAA 934
DR	P-PSDB; AAO15456.	Db	875 ACCTGGAAAGATAACCCCTAACCTGTAACCTGAACCTGAAAGAAA 934
PT	New isolated modulator of antigen receptor signalling protein or its fragments, useful for treating malignant disorders such as myeloid malignancies, autoimmune disorders and myeloproliferative disorders.	Qy	1054 GCTGAAAGCTCCCTCTGTTGAAAG---CTGCACAGGGGAGAGCTCTTCAG 1110
PT	Claim 10; FIG 1A; 110pp; English.	Db	935 GCTGAAAGCTCCCTCTGTTGAAAGACCTGCGTGGAGCTGGATGA 994
PS	The invention comprises the amino acid and coding sequences of modulator of antigen receptor signalling (MARS) proteins. The MARS protein is a putative tumour suppressor gene and exhibits structural and sequence similarity to the Scr-like adaptor protein (SAP). The MARS DNA and protein sequences of the invention are useful for the treatment of myeloid malignancies (e.g. acute myelogenous leukaemia), autoimmune disorders, immunosuppression, myeloproliferative disorders and malignancies related to the regulation of tyrosine kinases (e.g. breast cancer). The present cDNA sequence encodes a mouse MARS protein.	Qy	1111 TGAGGCTCTCCGGAGTCCTCACATGCCCTGAATGAGGAGACCTGGTCTCTTT 1170
XX	Sequence 1348 BP; 324 A; 365 C; 362 G; 277 T; 0 U; 0 Other;	Db	995 TGAGGCTCTCCGGAGTCCTCACATGCCCTGAATGAGGAGACCTGGTCTCTTGATGA 1054
XX	Query Match 47.5%; Score 561.4; DB 6; Length 1348;	Qy	1171 GGATGATGCCCTAG 1183
XX	Best Local Similarity 78.2%; Pkd. No. 7.1e-142;	Db	1055 TGCTTAGCCCTGG 1067
XX	Matches 714; Conservative 0; Mismatches 191; Indels 8; Gaps 3;	XX	RESULT 11
XX	274 AGGGCCCCAAAGCCTAACCTGTCACGCCAGAGCATGGCTCTTC 333	ID	AAS74748 standard; cDNA; 603 BP.
XX	160 AGGACCTGAAAGCCTGACTGTTGAGCTGAGTACCTCAT 217	XX	AAS74748;
XX	334 CAAGCCTTGTGACAACAAATTCTCCCTGATGATGCTCTGAGCTGAGG 393	XX	AC
XX	218 CAAACGTCATGGAAACCTTCCCTTCAGTTCAAGGCTGCTGCT 277	XX	DT
XX	394 AACATGGAAAGTCTGCCAGCAGAAATACTCTGCGAACGCCAAGCTTGAGTTCTC 453	XX	13-FBB-2002 (first entry)
XX	278 AGTGTGGAAAGTTGTCCAGCAGAGGAAACCT--CCAGGCCAACGCCAGCTCTC 334	XX	DNA encoding novel human diagnostic protein #10552.
XX	454 TGTCTAAGGCCAGGACACTGTGACATGCCAGZAGAGGAAAGGCCAGCCTGGGC 513	XX	XX Human; chromosome mapping; Gene mapping; Gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	335 TGGTCCAGCCAGAACCCGGTGTATGGAAAGAACAAAGGTTACAGCTGTGGC 394	XX	Homo sapiens.
XX	514 CCTGGGAGTTCCGGAGGTGCCCGAGCTGTGCTGAGCTGGGAGCCATT 573	XX	OS
XX	395 CCTGGGAGTTCCAGGGTGTACAGGAAAGACTATCTGAGCTGGGAGCCCT 454	XX	PN
XX	574 GACCATGCTCTGAGGATGAGACTGCTGTGTTGAGCTGGACGGCTGAGA 633	XX	W0200175057-A2.
XX	455 GACCATCATCTGAGGAGATGGAGATGGAGCTGTGAGCTGGAGCTGGAGA 514	XX	XX
XX	634 GTATACATCCCACCGTCACTGGGCAAAAGTCTCCATGCTGTATGAGGGCT 693	XX	11-OCT-2001.
XX	515 GTACACAGCCAGTGTATGGGTAAGTGGCTACAGGGCT 574	XX	XX
XX	694 GAGCAGGGAAAGGAGGAACCTGGACCTGGAACTCTGGAACTCTGGCTCT 753	XX	30-MAR-2001; 2001WO-US008631.
XX	575 GAGCCGGAAAGCGAGGAACACTCTGGTACCTGGAAACCCGGAAACGGCTCT 634	XX	XX
XX	754 CATCGGGAGCCAGCAGGAGGCTTACTCTCTGCTGCTCTGCTGCTCT 813	XX	31-MAR-2000; 2000US-00547217.
XX	635 CATCGGGAGAGCCAGCAGGAGGCTTACTCTCTGCTGCTCTGCTGCTCT 694	XX	PR
XX	814 TGCATCTGAGCCGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 873	XX	PA
XX	695 TGCATCTGAGCCGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 754	XX	(HYSEQ INC.
XX	874 CATCPACCGCCCTGAGGATACAGACATACGGATACAGACATACGGATACAG 933	XX	XX
XX	755 CATTCACACGCTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 914	XX	Drmanac RT, Liu C, Tang YT;
XX	Qy	XX	XX
XX	Db	WPI; 2001-639362/73.	
XX	Qy	P-PSDB; ABG10561.	
XX	Db	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.	
XX	Qy	Claim 1; SEQ ID NO 10552; 103PP; English.	
XX	Db	The invention relates to isolated polynucleotide (I) and polypeptide (II) and polypeptide chain (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligoners and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polynucleotide and polypeptide sequences have applications in	

diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS16197-AAS16564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this parent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct](ftp://wipo.int/pub/published_pct) sequences

P-PSDB; ABG10560  
New isolated poly-  
diagnostics, for  
responsibility for  
biodiversity.

Claim 1: SEQ ID N

P-PSDB; ABG10560.  
New isolated Polynucleotide and encoded polypeptides useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.  
Claim 1; SEQ ID NC\_10551; 103bp; English.

Qy	482	GAAGCAGAGGAAAGGCAAGCCGGCGCCCTGGGAGTTCCGGAGGTGGCCCG	541
Db	241	GAAGGAGAGAGAGAACGCCACAGCCGTCGGCGAGTTCGGAGGTGGCCCG	300
Qy	542	GCGAGCTGCGCTGAGACTGGGAGCCATTGACCATGCTCTGAG	589
Db	301	GCGAGCTGCGCTGAGACTGGGAGCCATTGACCATGCTCTGAG	348
OS	XX		
PN	XX		
RESULT 14			
ID	ABQ99151	standard; DNA; 875 BP.	
XX	XX		
AC	AC	ABQ99151;	
XX	XX	04-NOV-2002 (first entry)	
DT	DT		
XX	XX		
DB	Human ORF958 coding sequence.		
XX	XX		
(LEAC/)	LEACH M. D.	Cystostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;	
(MEHR/)	MEHRBAN P.	KW; Antinflammatory; Gene therapy; human; ORFX; atherogenic; Platelet;	
(CONL/)	CONLEY P. B.	human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;	
(TOPP/)	TOPPER J. N.	cancer; cardiovascular disease; allergy; autoimmune disease;	
(LAWD/)	LAW D.	wound healing; blood coagulation disorder; inflammatory disorder; ds.	
PA	XX		
PPI	XX		
Leach MD, Mehraban P, Conley PB, Topper JN, Law D;	XX		
WPI; 2002-625554/67.	XX		
P-PSDB; ABP64106.	XX		
XX	XX		
New polypeptide designated ORFX are present in human atherogenic cells	XX		
and are useful to prevent and treat ORFX-associated disorders including	XX		
cancer, allergy, wound healing or autoimmune, cardiovascular or	XX		
inflammatory disease.	XX		
Claim 2: SEQ ID NO 951; 78pp; English.	XX		
The present invention relates to novel human ORFX polypeptides and their	XX		
coding sequences (ABP6361 and ABQ98194-ABQ99267). The sequences	CC		
were discovered in human atherogenic cells, in particular in platelets	CC		
and human umbilical vein endothelial cells (HUVEC) and are expressed in	CC		
many other tissues as well. Atherogenic cells which have the	CC		
potential to develop atherosclerotic plaques. The ORFX polypeptides and	CC		
nucleic acids are useful for treating or preventing a pathological	CC		
condition associated with an ORFX-associated disorder, e.g. cancer,	CC		
cardiovascular disease, allergy, autoimmune disease, wound healing, blood	CC		
coagulation disorders or inflammatory disorders. Note: The sequence data	CC		
for this patent did not form part of the printed specification, but was	CC		
obtained in electronic format directly from the USPTO web site at	CC		
seqdata.uspto.gov/sequence.html?DocID=20020082206	CC		
XX	XX		
Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;	XX		
Score 348; DB 6; Length 444;	XX		
Query Match 29.4%; Pred. No. 3.7e-84;	XX		
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;	XX		
Matches 348; Conservative 0; N mismatches 0; N indels 0; N gaps 0;	XX		
Qy	242	CCTAGGACCAAGGA.CACTGGAGACTTCCAGAAGGGCCCCAAGGCCCTAACCTGTCAG	301
Ddb	1	CCTAGGACCAAGGA.CACTGGAGACTTCCAGAAGGGCCCCAAGGCCCTAACCTGTCAG	60
Qy	302	CCAGAGCATCGTCTGAGTGCTGAGTCAACCAATTTCCTCCAGAGCTTGTGACA	361
Ddb	61	CCAGAGCATCGTCTGAGTGCTGAGTCAACCAATTTCCTCCAGAGCTTGTGACA	120
Qy	362	TCTGATGATCTGCTCTGAGTGCTGAGTCAACCAATTGGAAAGTCGCCCAGCAAGAA	421
Ddb	121	TCTGATGATCTGCTCTGAGTGCTGAGTCAACCAATTGGAAAGTCGCCCAGCAAGAA	180
Qy	422	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	181	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	423	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	182	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	424	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	183	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	425	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	184	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	426	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	185	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	427	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	186	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	428	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	187	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	429	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	188	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	430	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	189	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	431	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	190	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	432	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	191	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	433	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	192	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	434	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	193	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	435	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	194	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	436	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	195	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	437	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	196	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	438	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	197	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	439	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	198	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	440	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	199	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	441	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	200	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	442	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	201	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	443	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	202	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	444	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	203	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	445	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	204	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	446	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	205	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	447	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	206	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	448	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	207	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	449	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	208	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	450	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	209	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	451	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	210	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	452	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	211	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	453	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	212	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	454	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	213	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	455	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	214	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	456	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	215	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	457	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	216	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	458	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	217	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	459	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	218	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	460	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	219	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	461	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	220	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	462	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	221	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	463	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	222	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	464	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	223	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	465	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	224	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	466	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	225	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	467	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	226	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	468	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	227	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	469	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	228	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	470	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	229	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	471	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	230	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	472	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	231	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	473	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	232	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	474	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	233	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	475	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	234	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	476	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	235	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	477	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	236	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	478	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	237	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	479	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	238	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	480	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	239	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	481	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	240	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	482	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	241	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	483	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	242	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	484	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	243	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	485	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	244	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	486	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	245	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	487	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	246	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	488	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	247	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	489	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	248	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	490	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	249	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	491	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	250	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	492	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	251	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	493	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	252	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	494	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	253	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	495	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	254	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	496	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	255	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	497	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	256	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	498	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	257	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	499	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	258	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	500	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	259	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240
Qy	501	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	481
Ddb	260	AAATCTCTGCAAGCCCAAGCTTGAGTCTCTGCTCCAGGCCAGGACCTGTGACCATG	240

Qy 843 GGATCCACTGGCTTGAATAAGCTGACATCTCACCCGCCCTACCTTCCTCAC 902  
 Db 2 GGATCCACTGGCTTGAATAAGCTGACATCTCACCCGCCCTACCTTCCTCAC 61  
 Qy 903 TCCAGGCCCTGGGACCATTACTCTGAGCTGGGATGACATCTGCTGCCTACTCAAGG 962  
 Db 62 TCCAGGCCCTGGGACCATTACTCTGAGCTGGGATGACATCTGCTGCCTACTCAAGG 121  
 Qy 963 AGCCCTGTGTGCTGCTGGCAGGGTGGCCGCCGCCCCGAGATAACCCCTACCTGGA 1022  
 Db 122 AGCCCTGTGTGCTGAGGGTGGCCGCCGCCCCGAGATAACCCCTACCTGGA 181  
 Qy 1023 CTGTGCAGAGGACACACTCAACTGGAAAGGCTGGACAGCTCCCTCGTTCTGAG 1082  
 Db 182 CTGTGCAGAGGACACACTCAACTGGAAAGGCTGGACAGCTCCCTCGTTCTGAG 241  
 Qy 1083 CTGCCACAGGGAGGAGCTCTTCTAGTAGGGGTTCGGGAGTCCTGAAGCTCTACA 1142  
 Db 242 CTGCCACAGGGAGGAGCTCTTCTAGTAGGGGTTCGGGAGTCCTGAAGCTCTACA 301  
 Qy 1143 TAGCCTGAATGAGCAGGGCTCTCTGGATGCTCTGGATGCTCTGGATGCTAG 1183  
 Db 302 TCAGCCTGAATGAGCAGGGCTCTCTGGATGCTCTGGATGCTCTGGATGCTAG 342

RESULT 15  
 AAS70181 ID AAS70181 standard; cDNA; 211 BP.  
 XX AC AAS70181;  
 XX DT 13-FEB-2002 (first entry)  
 XX DE DNA encoding novel human diagnostic protein #5985.  
 XX KW Human; chromosome mapping; gene therapy; forensic;  
 XX Food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX DR WPI: 2001-639362/73.  
 DR -PSDB; AB05994.

XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.  
 XX Claim 1: SEQ ID NO 5985; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics, forensics, gene mapping, identification of mutations in expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS6197-AAS91564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/patent\_pct\_sequences  
 CC Sequence 211 BP; 50 A; 51 C; 72 G; 38 T; 0 U; 0 Other;  
 SQ Score 191.2; DB 5; Length 211;  
 Query Match 16.2%; Score 98.5%; Pred. No. 9.6e-42;  
 Best Local Similarity 98.5%; Mismatches 0; Indels 0; Gaps 0;  
 Matches 193; Conservative 0;  
 YY 584 TCTGAGGATGGAGACTGTGGACGGTCTGAGTCTAGGGAGAGAAGTATAACATC 643  
 Db 16 TCCAAGGATGGAGACTGTGGACGGTCTGAGTCTAGGGAGAGAAGTATAACATC 75  
 XX 584 TCTGAGGATGGAGACTGTGGACGGTCTGAGTCTAGGGAGAGAAGTATAACATC 643  
 YY 644 CCCAGCGTCCACSTGGGAAAGTCTCCATGGGAAAGTCTCCATGGGAAAGTCTCCATGGGAG 703  
 Db 76 CCCAGCGTCCACSTGGGAAAGTCTCCATGGGAAAGTCTCCATGGGAAAGTCTCCATGGGAG 135  
 YY 704 AAGCAGAGGAACCTGCTGTGTACCTGGGACCTGGACCTGGCTCATCGGGAG 763  
 Db 136 AAGCAGAGGAACCTGCTGTGTACCTGGGACCTGGACCTGGCTCATCGGGAG 195  
 YY 764 AGCAGAGGAACCTGCTGTGTACCTGGGACCTGGACCTGGCTCATCGGGAG 779  
 Db 196 AGCAGAGGAACCTGCTGTGTACCTGGGACCTGGACCTGGCTCATCGGGAG 211

Search completed: February 19, 2004, 21:51:39  
 Job time : 1105 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:54 ; Search time 7832.77 Seconds

(without alignments)  
 4510.152 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 aaccttagctccaaaggacc.....tctttggatgtccatg 1183

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1491109276 residues

Total number of hits satisfying chosen parameters:

55026578

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: em\_estba:\*

2: em\_estbum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_ntc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_eston:\*

17: em\_ess\_hum:\*

18: em\_ess\_inv:\*

19: em\_ess\_pln:\*

20: em\_ess\_vrt:\*

21: em\_ess\_fut:\*

22: em\_ess\_mam:\*

23: em\_ess\_mus:\*

24: em\_ess\_pro:\*

25: em\_ess\_rid:\*

26: em\_ess\_phg:\*

27: em\_ess\_vrl:\*

28: gib\_ess1:\*

29: gib\_9992:\*

RESULT 1  
 AL541041 LOCUS AL541041 Homo sapiens PLACENTA Homo sapiens mRNA linear EST 12-MAY-2003  
 DEFINITION 5-PRIMB, mRNA sequence.  
 ACCESSION AL541041  
 VERSION AL541041\_2 GI:30544829  
 KEYWORDS EST  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens (human)  
 REFERENCE 1 (bases 1 to 1201)  
 AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polley,D.  
 TITLE Unpublished (2001)  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12871733.

#### SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	921.4	77.9	1201	9 AL541041	AL541041
2	860.2	72.7	1020	12 EQ054281	BQ54281 AGENCOURT
3	817	69.1	1002	12 EQ052308	BQ52308 AGENCOURT
4	731.4	61.8	1069	12 EQ052468	BQ52468 AGENCOURT

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

BP 191 9106 ERY ceder - France  
 Authors Li,W.B., Gruber,C., Jesse,J. and Polley,D.  
 Title Unpublished (2001)  
 Comment On Feb 15, 2001 this sequence version replaced gi:12871733.

#### ALIGNMENTS

BP 191 9106 ERY ceder - France  
 Email: segrégogénome.cns.fr , Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9825.r For more information about this cluster, see http://www.genoscope.cns.fr/  
 http://www.genoscope.cns.fr/cluster.cgi?seq=CS005.F12Q1&cluster=9825.r . Contact : Feng Liang Email : fliang@lifetech.com URL :

GATURES						
source	Query Match	Best Local Similarity	DB 9;	DB 10;	Length	Organism
Faraday Avenue Genoscope sequence ID : CSODE005AF12Q81.	1. .1201	97.9%	Score 921.4;	Pred. No. 1.5e-221;	1201;	"Homo sapiens"
	/mol type="mRNA"					
	/db_xref="taxon" value="606"					
	/clone="CSODE005IK23"					
	/tissue type="PLACENTA"					
	/clone lib="Homo sapiens PLACENTA"					
	/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a Not-I oligo(dT) primer. Five prime end enriched double strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."					
ORIGIN						
	Query Match	77.9%	Score 921.4;	DB 9;	Length 1201;	
	Best Local Matches	919;	Conservative	12;	Mismatches 8;	Indels 0;
						Gaps 0;
	4 TAGAGCTCCAAAGGACCCCAGCCCTGGCTGTGAGCAGAGCTCAAGGGCCCTGGGCCCT 63					
	65 TAGAGCTCCAAAGGACCCCAGCCCTGGCTGTGAGCAGAGCTCAAGGGCCCTGGGCCCT 124					
	64 TCCCCCTCTCTGGCTGGCTGGGGTTCAGTCAGATCCCTAAGGAGCAT 123					
	125 TCCCCCTCTGGCTGGCTGGCTGGGGTTCAGTCAGATCCCTAAGGAGCAT 184					
	124 GGGGAAGCTGATCCATCCCTGGTGTACAACCTGCTGACTCGAGAACAGATGCTGACCTACC 183					
	185 GGGGAAGCTGATCCATCCCTGGTGTACAACCTGCTGACTCGAGAACAGATGCTGACCTACC 244					
	184 CAAACAAACACCTAGCCCTCCCTGAGAATGCTCCAGGGCTGAGAGAGTCTGGGTGTGCC 243					
	245 CAAACAAACACCTAGCCCTCCCTGAGAATGCTCCAGGGCTGAGAGATCTGGGTGTGCC 304					
	244 TAGGACCAAGGACATGGCAACTTCAGAGGGCCCCAAAGCCCTAACCTGTCAAGGCC 303					
	305 TAGGACCAAGGACATGGCAACTTCAGAGGGCCCCAAAGCCCTAACCTGTCAAGGCC 364					
	304 AGAGGATGCTCTCGAGCTGAGCTGTTCCAAAGGCTTGTGATAAAACCAATTCCCTC 363					
	365 AGAGGATGCTCTCGAGCTGAGCTGTTCCAAAGGCTTGTGATAAAACCAATTCCCTC 424					
	364 GATGATGTCGTTCTGAGTGCCTGCTGGTGGAGAACATGGAAAGTCCTCCAGAGGAAA 423					
	425 GATGATGTCGTTCTGAGTGCCTGCTGGTGGAGAACATGGAAAGTCCTCCAGAGGAAA 484					
	424 ATCTTGCCAAAGGCCAAGCTGCTGGCTCTGAGTCCTCTGCAAGGGACCTGACATGG 483					
	485 ATCTTGCCAAAGGCCAAGCTGCTGGCTCTGAGTCCTCTGCAAGGGACCTGACATGG 544					
	484 AGCAGAGAAAGGAGGCCAACAGCCTGGGCGCTGGGCAAGTTCCGGCAAGTGGCCGGC 543					
	545 AGCAGAGAAAGGAGGCCAACAGCCTGGGCGCTGGGCAAGTTCCGGCAAGTGGCCGGC 604					
	544 CGAGCTGTCCTGAGACTCGGGAGGCCATTGACCATGCTCTGAGATGGAGCTGGTG 603					
	605 CGAGCTGTCCTGAGACTCGGGAGGCCATTGACCATGCTCTGAGATGGAGCTGGTG 664					
	604 GACGGCTGTCGAGTCTAGGCGAGACTATTAATCCCAGGCTCAAGCTGGGCAA 663					
	665 GACGGCTGTCGAGTCTAGGCGAGACTATTAATCCCAGGCTCAAGCTGGGCAA 724					
	664 AGTCTCCATGGGCTGTATGAGGGCTGAGGAGAACATGGAGCTGCTGGT 723					
	725 AGTCTCCATGGGCTGTAGGAGGGCTGAGGAGAACATGGAGCTGCTGGT 784					
	724 GTTACTCTGGGAACTCTGGGAGGCCCTCTGGAGGCCCTCTGGAGGCCCTCTGGT 783					
	785 GTTACTCTGGGAACTCTGGGAGGCCCTCTGGAGGCCCTCTGGAGGCCCTCTGGT 844					

RESULT 2									
QY	784	TTRACTCTCTGTAGTGGCTCAATGGTGTGATCTACCGCCCTGCATGCTGGACGCCATGACATACAG	843						
Db	845	TTRACTCTCTGTAGTGGCTCAATGGTGTGATCTACCGCCCTGCATGCTGGACGCCATGACATAMAG	904						
QY	844	GATCCACTGCGCTGAATGGTGTGATCTACCGCCCTGCATGCTGGACGCCATGACATCAG	903						
Db	905	GATCCACTGCGCTGAATGGTGTGATCTACCGCCCTGCATGCTGGACGCCATGACATTCAG	964						
QY	904	CGAGGCCCTGGTGGACATTACTCTGGCTGGATGA	942						
Db	965	CGAGGCCCTGGTGGACATTACTCTGGCTGGATGA	1003						
<b>DEFINITION</b>									
LOCUS	BQ054281	BQ054281	1020 bp	mRNA	linear	EST	29-MAR-2002		
DEFINITION	AGENCOURT	6830234	NTH_MGC_106	Homo sapiens	cdNA clone	IMAGE:5936362			
5'-tRNA sequence.	BQ054281								
ACCESSION	BQ054281.1								
VERSION	BEST								
KEYWORDS	EST.								
SOURCE	Homo sapiens								
ORGANISM	Homo sapiens								
EUROPORT	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.									
REFERENCE	1 (bases 1 to 1020)								
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/							
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)								
JOURNAL	Unpublished (1999)								
COMMENT	Contact: Robert Strausberg, Ph.D.								
Email:	cgapbs-r@mail.nih.gov								
Tissue Procurement:	Dr. Daniel McVicar, DBS/NCI								
CDNA Library Preparation:	Rubin Laboratory								
CDNA Library Arrayed by:	The T.M.A.G.E. Consortium (LLNL)								
DNA Sequencing by:	Agencourt Bioscience Corporation								
Clone distribution:	MGC clone distribution information can be found through the T.M.A.G.E. Consortium/LLNL at:								
	http://image.llnl.gov								
Plate:	LICM2125	row: j	column: 11						
High quality sequence stop:	556.								
FEATURES	Location/Qualifiers								
	1..1020								
source	/organism="Homo sapiens"								
	/mol type="mRNA"								
	/db_xref="taxon:9605"								
	/clones="IMAGE:5936363"								
	/tissue type="natural killer cells, cell line"								
	/lab_host="DH10B (ophage-resistant)"								
	/clones="NIH_MGC_106"								
	/note="Organ: blood; Vector: P0TB7; Site 1: XbaI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into SCoRI/XbaI sites using the following 5' dlapcpr: GccACGAG (G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and SuperScript II RT (Life Technologies). Note: this is a NTH_MGC Library."								
ORIGIN									
Query Match	72.7%	Score 860.2;	DB 12;	Length 1020;					
Best Local Similarity	96.1%	Pred. No. 3	9e-206;						
Matches	956;	Conservative	0;	Mismatches					
					Indels	9;	Gaps	7;	
QY	187	ACCAACCTAGCTCTCCCTGAAGATCCTCCAGGTGAGAGCTCTGGGTGTCCTAG	246						
Db	1	ACCAACCTAGCTCTCCCTGAAGATCCTCCAGGTGAGAGCTCTGGGTGTCCTAG	60						
QY	247	GACCGAGGACACTGGCAGACTTGAGAGCTTCAAGCTTGTGACAAACGCTTAC	306						
Db	61	GACCAAGGACACTGGCAGACTTGAGAGCTTCAAGCTTGTGACAAACGCTTAC	120						
QY	307	GCATGGCTCACAGAACGCTTCCAAAGCTTGTGACAAACGCTTAC	366						





REFERENCE NIH-MGC http://mgc.ncbi.nlm.nih.gov/	Do	539 CCATCGTCTCTGAGGATGGAGACTGGTGCTGAGCTGGAGGT 598
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC) Unpublished COMMENT Contact Robert Strausberg, Ph.D. Email: cgsaps@MAIL.nih.gov Tissue Procurement: DCPD/DTP cDNA Library Preparation: Life Technologies, Inc. DNA Arrayed by: The I.M.A.G.E. Consortium (LNU)	Qy	636 ATTACATCCACCGTCACCTGGCAAGAGTCTCCATTGGCTGA 695
DNA Sequencing by: Incyte Genomics, Inc.	Db	599 ATTACATCCACCGTCACCTGGCAAGAGTCTCCATTGGCTGA 658
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNU at: <a href="http://image.lnl.gov/IMAGE/LNL1082">http://image.lnl.gov/IMAGE/LNL1082</a> row: i column: 01	Qy	696 GCAGGGAGA-AAGCAGAGAACTGGTACCTGGAAACCTGGAGGT 752
High quality sequence stop: 657.	Db	659 GCAGGGAGAACAGCAGAGAACTGGTACCTGGAAACCTGGAGGT 718
FEATURES source	Do	753 TCATCCGGAGAGCCAGACCCAGGAGGGCTTTA 786
Location.Qualifiers 1. 778	Do	719 CATCCCGGAAAGCAGACCCAGGAGGGCTCTTA 752
1. /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4429896" /tissue_type="denocarcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_id="NIH-MGC 91" /note="Organ: prostate; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."	Qy	RESULT 6
ORIGIN	Db	B0054265 LOCUS B0054265 DEFINITION AGENCOURT-6830248 NIH_MGC_106 Homo sapiens cdNA clone IMAGE:5936339
COMMENT	Qy	/db_xref="taxon:9606" /version B0054265.1 GI:19813605 /keywords EST /organism Homo sapiens (human) /host_molecule Homo sapiens (human) /mammal Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. Unpublished (1999)
COMMENT	Db	REFERENCE NIH-MGC http://mgc.ncbi.nlm.nih.gov/ AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC) TITLE Unpublished JOURNAL Unpublished COMMENT Contact: Robert Strausberg, Ph.D. Email: cgsaps@mail.nih.gov Tissue Procurement: Dr. Daniel McVicar, DBS/NCI cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov/IMAGE/LNL1082">http://image.lnl.gov/IMAGE/LNL1082</a> row: i column: 12 Place: LNL2125 High quality sequence stop: 515. Location/Qualifiers 1. .986 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5936339" /tissue_type="natural killer cells, cell line" /lab_host="DH10B (phage-resistant)" /clone_id="NIH-MGC 106" /note="Organ: blood; Vector: PDB7; Site_1: XbaI; Site_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGCAAGAG (G) Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and SuperScript II RT (Life Technologies). Note: this is a NIH-MGC Library."
Query Match Best Local Similarity 96.9%; Score 673.6; DB 12; Length 778; Matches 730; Conservative 0; Mismatches 19; Indels 5; Gaps 4;	Qy	36 GTGACAGCTCAAAGGCCCTGGCCTCCCTGGCTGGTGGAGGT 95
Db	1 GTGACAGCTCAAAGGCCCTGGCCTCCCTGGCTGGTGGAGGT 59	
Qy	96 TCCCAGTCCAAATCCTCTTCTTCTGATCATCCCTGTGTACAATT 155	
Db	60 TCCCAGTCCAAATCCTCTTCTGATCATCCCTGTGTACAATT 119	
Qy	156 GETGACTGAGACAGATGCTGAGTACCCAAACCTAGCCTCCCTGAGAGATC 215	
Db	120 GETGACTGAGACAGATGCTGAGTACCCAAACCTAGCCTCCCTGAGATC 179	
Qy	216 TCCCAAGCTGAGAGTTCTGGTCTCTAGGACCAAGAACCTGGCAGACTTCCTGAGAG 275	
Db	180 TCCAGCTGAGAGTTCTGGTCTCTAGGACCAAGAACCTGGCAGACTTCCTGAGAG 239	
Qy	276 GGCCCCAAAACCCCTAACCTTCCAGCCAGGCAAGCTCTCACAGAGCTTCCA 335	
Db	240 GGCCCCAAAACCCCTAACCTTCCAGCCAGGCAAGCTCTCACAGAGCTTCCA 299	
Qy	336 AGCCCTTGATGACAAACCAATTCCCTCGATGATGTGCTCTGAGTGGCTGCTGAGAA 395	
Db	300 AGCCCTTGATGACAAACCAATTCCCTCGATGATGTGCTCTGAGTGGCTGAGAA 359	
Qy	396 CAATGGAAAGTGGCCAGGACCTGACCATGGAAAGAACCTGAGTGGCTCTGAG 455	
Db	360 CAATGGAAAGTGGCCAGGACCTGACCATGGAAAGAACCTGAGTGGCTCTGAG 419	
Qy	456 TCCAAAGCCAGGACCTGACCATGGAAAGAACCTGAGTGGCTCTGAG 515	
Db	420 TCCAAAGCCAGGACCTGACCATGGAAAGAACCTGAGTGGCTCTGAG 478	
Qy	516 TGGCAGTTCCGCAAGGTCCTCCCGCCGCTGAGACTGGGAGCCATCA 575	
Db	479 TGGCAGTTCCGCAAGGTCCTCCCGCCGCTGAGACTGGGAGCCATCA 538	
Qy	576 CCATCGTCTCTGAGGATGGAAACTGGTGGACGGTGTCTGAGAAGAGAGT 635	
Db	61 TACCCAAACCAACCTAAGCTCCCTGAGATCCTCCAGGCTGAGAAGAGTCTGGGT 120	

		FEATURES	Location/Qualifiers
Qy	240	GTCCTAGGACCAAGGACAACCTGGCAGACTTCAGBAGGCCAAAGCCCTAACCTGTCC	299
Db	121	GTCCTAGGACCAAGGACAACCTGGCAGACTTCAGBAGGCCAAAGCCCTAACCTGTCC	180
Qy	300	AGCCAGAGCATGCGTCTCAGAGCTCTTCAGCAAGCCTTGATGACAACCAATTTC	359
Db	181	AGCCAGAGCATGCGTCTCAGAGCTCTTCAGCAAGCCTTGATGACAACCAATTTC	240
Qy	360	CCTCATGATGCTGTTCTAGTGTCTGAGGAGTGTGCCAGGAGCAGAAGCTGGAGGAA	419
Db	241	CCTCATGATGCTGTTCTGAGGAGTGTGCCAGGAGCAGAAGCTGGAGGAA	300
Qy	420	GAAATCTCTGCAAGCCAAGCTGTGTTGCAAGCCAGGAACCTGTGACCA	479
Db	301	GAAATCTCTGCAAGCCAAGCTGTGTTGCAAGCCAGGAACCTGTGACCA	360
Qy	480	TGGAGCAGGAGGAGCANGCCAGGCTGGCCAGCTGGAGCTTTCGGAGGTGGCC	539
Db	361	TGGAGCAGGAGGAGCANGCCAGGCTGGCCAGCTGGAGCTTTCGGAGGTGGCC	420
Qy	540	CGGCCAGCTGTCCTGAGACTGGGAGAACCATGACCATGCTCTCTGAGGATGGAGCT	599
Db	421	CGGCCAGCTGTCCTGAGACTGGGAGAACCATGACCATGCTCTCTGAGGATGGAGCT	480
Qy	600	GTTGACCTGCTGCTGCTGTAAGTCTCAGGCAAGGAGTATAACATCCCAGCAGCTGG	659
Db	481	GTTGACCTGCTGCTGTAAGTCTCAGGCAAGGAGTATAACATCCCAGCAGCTGG	540
Qy	660	GCAAAGTCTCCATGGTGGCTGTATGAGGGCTCTGAGCAGGGAAAGCAGGAAACTGC	719
Db	541	GCAAAGTCTCCATGGTGGCTGTATGAGGGCTCTGAGCAGGGAAAGCAGGAAACTGC	600
Qy	720	TGTGTCTACTGGAAACCCTGGGGGG-CCTTCTCTCATCGGGAG-AGCCAACCGGG	777
Db	601	TGTGTCTACTGGAAACCCTGGGGGGCTCTCATCGGGAG-AGCCAACCGGG	660
Qy	778	AGGTCTTACTCTCTGTCAAGCT-GCCG-CCTGTGATCCCTGG-ACCGATCA	833
Db	661	AGGTCTTACTCTCTGTGATCCCTGG-ACCGATCA	720
Qy	834	GACATACAGG 844	
Db	721	GACCTTACAGG 731	
RESULT	7		
LOCUS	BQ03486	878 bp	mRNA linear EST 29-MAR-2002
DEFINITION	AGENCOURT 6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253		
ACCESSION	5' mRNA sequence.		
VERSION	BQ03486.1		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
MATERIAL	Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 878)		
AUTHORS	NIH-MGC http://mgc.ncbi.nlm.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
COMMENT	Contact: Robert Straubberg, Ph.D.		
COMMENT	Email: cgbps@mail.nih.gov		
TISSUE	Procurement: Dr. Daniel McVicar, DBS/NCI		
CDNA	Library Preparation: Rubin Laboratory		
CDNA	Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)		
DNA	Sequencing by: Agencourt Bioscience Corporation		
Clone	Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:		
	http://image.llnl.gov		
Plate	LUCM2122 Row: 1 column: 06		
	High quality sequence stop: 394.		
RESULT	8		

ACCESSION	AK088672	DEFINITION	Mus musculus 2 days neonate thymus mRNA linear	HTC 20-SEP-2003
KEYWORDS	ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.	FULL-LENGTH ENRICHED LIBRARY CLONE	E130022D24	PRODUCT: MODULATOR OF
ORGANISM	Mus musculus (house mouse)			
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	AUTHORS		
TITLE	High efficiency full-length cDNA cloning	JOURNAL		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	MEDLINE		
PUBLISHED	10349636	REFERENCE		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	AUTHORS		
TITLE	Normalisation and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	JOURNAL		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	MEDLINE		
PUBLISHED	20493974	REFERENCE		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaeiki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kirsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamoto, S., Hazama, M., Nishimura, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	AUTHORS		
TITLE	RIKEN integrated sequence analysis (RSA) system--3D4-format sequencing pipeline with 384 multicapillary sequencer	JOURNAL		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	MEDLINE		
PUBLISHED	20530913	REFERENCE		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	AUTHORS		
TITLE	Functional annotation of a full-length mouse cDNA collection	JOURNAL		
REFERENCE	Nature 409, 685-690 (2001)	REFERENCE		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	AUTHORS		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	JOURNAL		
JOURNAL	Nature 420, 563-573 (2002)	REFERENCE		
REFERENCE	6 (bases 1 to 2637)	AUTHORS		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, T., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramatsu, K., Hirooka, T., Hirozane, T., Katoh, H., Kawai, J., Kohjima, Y., Kondo, S., Konno, H., Kouada, M., Koya, S., Kuriiwara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sasaki, K., Sakazume, N., Sano, H., Sasaki, D., Shibusawa, K., Shinagawa, A., Shiraki, T., Scogbe, Y., Tasami, M., Tagami, A., Takayashi, F., Takami-Akainira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyo, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	JOURNAL		
TITLE	Direct Submission	REFERENCE		
JOURNAL	(16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-rs@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/), Tel: 81-45-903-9222, Fax: 81-45-503-9916	COMMENT		
CDNA library was prepared and sequenced in Mouse. Genome Encyclopedia Project of Genome Exploration Research Group in Riken				



prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

**JOURNAL** Genome Res.  
**AUTHORS** Konno,H., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,  
Kondo,Y., Akiyama,J., Nishi,K., Kitamura,T., Tashiro,H., Ichim,M.,  
Sumi,N., Iishi,Y., Nakamura,M., Hazama,M., Nishimi,T., Harada,A.,  
Yamamoto,R., Matsunaga,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanuki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsubara,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.  
**TITLE** RIKEN integrated sequence analysis (RISEA) system-384-format sequencing pipeline with 384 multi-capillary sequencer  
**JOURNAL** Genome Res.  
**PUBMED** 10530913  
**PMID** 11076861  
**REFERENCE** 4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

**JOURNAL** JOURNAL OF HEREDITY  
**AUTHORS** The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

**JOURNAL** JOURNAL OF HEREDITY  
**AUTHORS** Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carrinici,P., Fukuda,S., Furuno,M., Hanaguri,T., Hara,A., Hasinzu,W., Hayashida,K., Hayatsu,H., Hiramoto,K., Hiracka,T., Hirozane,T., Hori,F., Imotani,K., Iishi,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katch,H., Kawai,J., Kondo,S., Kondo,Y., Kojima,Y., Koura,H., Koya,S., Kuribayashi,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numaraki,R., Ohno,M., Onsato,N., Okazaki,Y., Saito,R., Seijita,H., Sakai,C., Sakaue,M., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Takanashi,F., Takaku-Akamira,S., Takechi,Y., Taraka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
**TITLE** Direct Submission  
**JOURNAL** JOURNAL OF HEREDITY  
**AUTHORS** Submitted (16-JULY-2001). Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, Fax: 81-45-503-9216)  
**COMMENT** cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
**FEATURES** URL:<http://genome.gsc.riken.go.jp/>  
URL:<http://Fantom3c.riken.go.jp/>  
**Source** 1. organism="Mus musculus"  
/mo\_type="mrna"  
/strain="C57BL/6J"  
/ch\_xref="C57BL/6J"

Sequence Comparison Results						
	Query	Match	Score	Length	Start	End
100% ID, 100% length, match=777, putative"						
/codon_start=1						
/protein_id="BAC27168.1"						
/db_xref="NCI:26326899"						
GB#QAR1SLRQEPLITIISNGLSLSEGRKSSSPSSPSSGPIQDPEVNSQEPHRKVTAVALGSFPA						
AEELFLPLPGGGAAFLIRESOOTRRYYSLSVRLSRPASWRIRHRYRQDNGNLYIT						
PRUTPSLHLWVHSELAGICCPDREPVLQRQLPDKDTPPPVTVTSSLWKK						
LDRSLLFLEAPSGEASLISSEGRLRESLSSYTLAEDPDLDA"						
ORIGIN						
Query Match	45.5%	Score 538.6 ; DB 11;	Pred. No. 1-5e-124;	Length 2974;		
Best Local Similarity	80.9%		Indels 6;	Gaps 2;		
Matches 653; Conservative Matches 0; Mismatches 148;						
Qy 380 GngctrcgtGAGGAATGGAACTGGCAAGCTGCCAGGAAGAAATTCTGCAAGGCCA 439						
Db 214 GTGAGCCTCTGTGAGTGGAAAGCTGTGACCTGGAAAGGAAAC--CTTCAGGCC 270						
Qy 440 AGTTTGAGTCTCCCTCTGTCAGGCCAACGGACCTGTGACCATGGAAAGCAAG 499						
Db 211 AGCCCCAGCTCTCTGTCCAGAACAGAAACCCTGTGTCATGCAAAGACAAAG 330						
Qy 500 GCGAACACGCTGCGCCCTGGCGAGGTTCCAGGCCAACGGACCTGGAAAGCTGGAGA 559						
Db 331 GTCAACACTGAGCCCTGGAGGTTCCAGGGTAACAGCCAGACTATCTCTGAGA 390						
Qy 560 CTCGGGGAGGCAATTGACATCTCTCTGAGGATGGAGACTGGTGAACGGTGTGCTCAA 619						
Db 391 CTGGGGAGGCGCTGACATCTCTGAGGATGGAGACTGGTGAACGGTGTGCTCAA 450						
Qy 620 GTCCTAGGGAGGATAAGATCCCAGCTGCCAGGCAAACTCTCCATGGTGG 679						
Db 451 GTCCTAGGGAGGATAAGATCCCAGCTGCCAGGCAAACTCTCCATGGTGG 510						
Qy 680 CTGTATAAGGGCTGAGCGGGAGAAGACAGAGGAACTGCTGTTGTAACCTGGAAACCT 739						
Db 511 CTGTACCGAGGGCTGAGCGGGAGAAGACAGGAAACTACTCTGTAACTGGAAACCC 570						
Qy 740 GGAGGGCCCTCTCATCGGGAGAGGCCAGGAGGCCAGGAGGCCCTTAACTCTGTGAGTC 799						
Db 571 GGAGGGCCCTCTCATCGGGAGGCCAGGAGGCCCTTAACTCTGTGAGTC 630						
Qy 800 CCCTCTAGCCCTCTGATCTGGGACGGATCAGACTACAGGATCCACTGCTTGTGAC 859						
Db 631 CGACTCTGCGCTGATCTGGACGGATCAGACTACAGGATACAGCTGTGAC 690						
Qy 860 AATGGGCTGGGNTACATCTACCGGCTCTCACCTCCCTCACTCCGGOCTGGTGCAC 919						
Db 691 AATGGGCTGGGNTACATCACTCCCTACCTCCCTACCTCCACCTGTGAG 750						
Qy 920 CATTACTCTGAGCTGGGTGAGCATCTGTGCTACTCTAAGGGCCCTCTGTCTGCAAG 979						
Db 751 CATTACTCTGAGCTGGGTGAGCATCTGTGCTACTCTAAGGGCCCTCTGTCTGCAAG 810						
Qy 980 AGGGCTGGGCTCTGGAAAGGATATACTGGCAAGCTGCTACTCTGTGACTCTGTGCTGAGGCCA 1039						
Db 811 AAAGCTTGGCCACTACTCTGGAAAGATACACCTCCACCTGTGCAACATCATCA 870						
Qy 1040 CTCAACTGGAAAGGTGGAGCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTGAG --CTGCAACAGGGAG 1096						
Db 871 CTAATTTGGAAAAGTTGGACCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTGAG 930						
Qy 1097 GAGTCCTCTGCTGAGCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTGAG 1156						
Db 931 GCATCTGCTGCTAGCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTGAG 990						
Qy 1157 GAGGTCTCTCTGCTGAGCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTGAG 1183						
Db 921 GCGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTGAG 1221						

RESULT 11	AK020837	926 bp	mRNA linear	HTC 20-SEP-2003
LOCUS	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone: A30009E21 Product: MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.			
DEFINITION				
ACCESSION	AK020837			
VERSION	1			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus: Mus musculus			
REFERENCE	Carninci, P., and Hayashizaki, Y.			
AUTHORS	Itoh,M., Konno,H., Okamura,S., Sugahara,Y., Shibata,K., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneida,Y., Ishikawa,T., Ozawa,M., Kanaka,T., Matsunatsu,M., Okazaki,Y., Muranatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042119			
REFERENCE	Shibata,K., Itoh,M., Aizawa,K., NagaoKA,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Sumi,N., Iishi,Y., Nakamura,S., Hazama,M., Nishime,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneida,Y., Ishikawa,T., Ozawa,M., Kanaka,T., Matsunatsu,M., Okazaki,Y., Muranatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.			
AUTHORS	Itoh,M., Konno,H., Okamura,S., Sugahara,Y., Shibata,K., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneida,Y., Ishikawa,T., Ozawa,M., Kanaka,T., Matsunatsu,M., Okazaki,Y., Muranatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multicapillary sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	20530913			
PUBMED	11076681			
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
AUTHORS	Functional annotation of a full-length mouse cDNA collection			
TITLE	Nature 409, 685-690 (2001)			
JOURNAL				
REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
AUTHORS	Analysis of the mouse transcriptome based on functional annotation of 60,700 full-length cDNAs			
TITLE	Nature 420, 563-573 (2002)			
JOURNAL				
REFERENCE	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,N., Hiramatsu,K., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramatsu,K., HiraoKA,T., HorI,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasuya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nihei,K., Nomura,K., Nummzaki,R., Ohno,M., Okazaki,Y., Okido,T., Ota,C., Saito,H., Saito,R., Sakai,C., Sahai,K., Sano,H., Sasada,D., Shibusawa,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasumishi,A., Yoshida,K., Yoshino,M., Muranatsu,M. and Hayashizaki,Y.			
AUTHORS	Direct Submission			
TITLE	Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-reg@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-303-9222, Fax: 81-45-503-9216)			
JOURNAL	Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for			
COMMENT				



Query Match Similarity	38.2%	Score	452.2	DB	13	Length	660	
Best Local Similarity	83.4%	Pred. No.	3_9e-103					
Matches	533; Conservative	0; Mismatches	103;	Indels	4;	Gaps	2;	
Db	4	GGGAGTTCGCCAGGTGCCGGCCAGCTGCGACTCGGGGCCATTGAC	576	FEATURES	Source			
Qy	517	GGGAGTTCGCCAGGTGCCGGCCAGCTGCGACTCGGGGCCATTGAC	576			High quality sequence start:	2	
Db	577	CATGTCCTGAGATGAGACTGGACGGCTGCTGAAGTCAGGAGAGATA	636			High quality sequence stop:	566.	
Qy	64	CATGTCCTGAGATGAGACTGGACGGCTGCTGAAGTCAGGAGAGATA	63	Location/Qualifiers				
Db	637	TAATCCCGGCAAGCTCCAGGGCRAACTCTCCATGGTGCTGTATAGGGCTGAG	696	1.	/organism="Homo sapiens"			
Qy	124	CCATGCCAGCTGATGGAGATGGTGCTGTATGGTGCTGTATAGGGCTGAG	183	/mol_type="mRNA"				
Db	697	CAGGGAAAGCAGAGGAACCTGTTACCTGGAACCTGGAGGGGCTTCCCAT	756	/clone_xref="IMAGE:4520382"				
Qy	184	CGGGAGAAGCCAGGAACTACCTGTTACCTGGAACCCGGCTTCCAT	243	/tissue_type="adenocarcinoma, cell line"				
Db	757	CGGGAGAAGCCAGGAACTACCTGTTACCTGGAACCCGGCTTCCAT	816	/lab_host="DH10B (phage-resistant)"				
Db	244	CGGGAGAAGCCAGGAACTACCTGTTACCTGGAACCCGGCTTCCAT	303	/clone_lib="NIH MGC 91"				
Qy	817	ATCTGGGACGGATCAGAACTACAGGATCCACTGCTGGCTGTACAT	876	/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;				
Db	304	ATCTGGGACGGATCAGAACTACAGGATACGGTGTGAACTGGCTGTACAT	363	Site 2: Sall; Cloned unidirectionally; Cligo dr primed.				
Qy	877	CTCACCGGCCTCACCTCCCTCACTCCGGCTGTGACATTACTTGAGTGGC	936	Average insert size 1.4 kb. Library enriched for				
Db	364	CTCACCTCCGCTCACCTCCCTCACTCCGGCTGTGACATTACTTGAGTGGC	423	full-length clones and constructed by Life Technologies.				
Qy	937	GGATGACATGGCTGCTGCTACTCAAGGGCCCTGGCTGAGGGCTGGCGTCC	996	Note: this is a NIH_NGC Library.				
Db	424	AGATGGCATGGTGTGTCCTGGGACGCCGCTGGTCTGGAGAAGTTGGCCACTACC	483					
Qy	997	TGGCAGGATAATACCCTTACCTGTGACTGTGAGGGACCACTCACTGAAAGGT	1056					
Db	484	TGGCAGGATAATACCCTTACCTGTGACTGTGAGGGACCACTCACTGAAAGGT	543					
Qy	1057	GGACAGCTCCCTCTGTGAGAAGCTGGGAGGAGTCTCTCTCAGTGA	1113					
Db	544	GGACCGTAGGCTCTGTGTTGGAGCACCTGAAAGTGGGGCATCTGCTCATTTGA	603					
Qy	1114	GGGTCTCGGGAGTCCCTCACTACATCAGCTCTGAGTACATCAGCTCTGAGA	1158					
Db	604	GGGCTCCAGAGT-CCTCACCTCTACCTCATCCGGTTGATGA	647					
RESULT 13	BG284179	566 bp mRNA	linear	EST	21-FEB-2001			
LOCUS	60240826F1	NIH_MGC_91	Homo sapiens	CDNA clone IMAGE:4520382	5			
DEFINITION								
ACCESSION	BC284179	mRNA sequence.						
VERSION	BC284179.1	GI:13034866						
SOURCE								
ORGANISM	Homo sapiens	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.						
COMMENT	1 (bases 1 to 566)	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)						
REFERENCE		Contact: Robert Strausberg, Ph.D.						
AUTHORS		Email: cgabbs@mail.nih.gov						
JOURNAL		Tissue Procurement: DCTD/DTP						
COMMENT		CDNA Library Preparation: Life Technologies, Inc.						
		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)						
		Clone distribution: MGC clone distribution information can be						

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10418 row: c column: 07

High quality sequence start: 2

High quality sequence stop: 566.

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

ORIGIN

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Site 2: Sall; Cloned unidirectionally; Cligo dr primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_NGC Library."

Query Match

34.2% Score 405; DB 12; Length 566;

Best Local Similarity 100.0%; Pred. No. 3e-91; Indels 0; Gaps 0;

Mismatches 0;保守性 0; Matches 0;

Indels 0; Gaps 0;

Location/Qualifiers

1.

/organism="Homo sapiens"

/mol\_type="mRNA"

/clone\_xref="IMAGE:4520382"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: Not I;

Email: humquery@sanger.ac.uk

Sanger Centre name : scdi0818.154136A

Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100 000 cDNA clones derived from pool-YT.lib\_NSPD cDNA library. Further information can be found at <http://www.sanger.ac.uk/Teams/Team69/>.

## FEATURES

## source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="20"  
 /clone\_lib="pool\_YT\_lib\_V\_SPD"

## ORIGIN

Query Match 27.3%; Score 322.4; DB 9; Length 597;

Best Local Similarity 99.7%; Fred. No. 2.3e-70;  
 Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 860 ATGGCTGCTGTACATCTCACCCGCCCTCACTCCCTCTGTGAC 919

Db 13 ATGGCTGCTGTACATCTCACCCGCCCTCACTCCCTCTGTGAC 72

Qy 920 CATTACTCTGAGCTGGGATGACATCTGCCTACTCAGGGCCCTGTGTCCTGAG 979

Db 73 CATTACTCTGAGCTGGGATGACATCTGCCTACTCAGGGCCCTGTGTCCTGAG 132

Qy 980 AGGGCTGGCCCGCTCCCTGGAAAGATAACCCCTAACCTGACTGTGAGAGGACACCA 1039

Db 133 AGGGCTGGCCCGCTCCCTGGAAAGATAACCCCTAACCTGACTGTGAGAGGACACCA 192

Qy 1040 CTCAACTGGAAAGAGGTGGAGAGCTGGCTCCCTCTGTGTTCTGAAGCTGCCACAGGGAGAG 1099

Db 193 CTCAACTGGAAAGAGGTGGAGAGCTGGCTCCCTCTGTGTTCTGAAGCTGCCACAGGGAGAG 252

Qy 1100 TCTCTCTCTGAGTGGGTCTGGAAAGTCCCTCAGGTTCTCATAGCTGTAATGACAG 1159

Db 253 TCTCTCTCTGAGTGGGTCTGGAAAGTCCCTCAGGTTCTCATAGCTGTAATGACAG 312

Qy 1160 GCTGTCTCTGGAGGTCTGGATGATGCTAG 1183

Db 313 GCTGTCTCTGGATGATGCTAG 336

## RESULT

BU944126

LOCUS BU944126 960 bp mRNA linear EST 18-OCT-2002  
 DEFINITION AGENCOURT 10545003 NIH\_MGC\_107 Homo sapiens cDNA clone  
 IMAGE:678350 5', mRNA sequence.

VERSION BU944126

KEYWORDS EST

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

COMMENT Contact: Robert Strausberg, Ph.D.

Email: rgs@sra.nih.gov

Tissue Procurement: ARCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The J.M.A.G.E. Consortium (JNIN)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the J.M.A.G.E. Consortium/JNIN at:  
<http://image.jnlni.gov>

Plate: LCM3049 row: m column: 21

High quality sequence stop: 628.

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6728350"

/issue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_107"

/clone="breast"; Vector: pORB7; Site: 1; ECOLI:

Site: 2; XbaI; cDNA made by Oligo-dT priming.

Directionally cloned into ECO RI sites using the following 5' adaptor: GCCACCGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library.

## ORIGIN

Query Match 27.1%; Score 321; DB 13; Length 960;

Best Local Similarity 88.4%; Pred. No. 7.1e-70;  
 Matches 381; Conservative 0; Mismatches 0; Indels 50; Gaps 1;

Qy 753 TCATCCGGAGAGGCCAGAGCCAGGAGGGCTTACTCTCTGTCAGCCGCC 812

Db 1 TCATCCGGAGAGGCCAGAGCCAGGAGGGCTTACTCTCTGTCAGCCGCC 60

Qy 813 CTGGCATCTGGACCCGGATGACACTACAGGATCCACTGCTGAGCTGCTGT 872

Db 61 CTGGCATCTGGACCCGGATGACACTACAGGATCCACTGCTGAGCTGCTGT 120

Qy 873 ACATCTAACGGGGCTCACTCCCTACTCCAGGCTCTGGACCATTACTCTGAGC 932

Db 121 ACATCTAACGGGGCTCACTCCCTACTCCAGGCTCTGGACCATTACTCTGAGC 176

Qy 933 TGGGGATGACATCTGGCTACTAAGGCCCCCTGTGCTGAGAGGCTGGCCGC 992

Db 177 -----

Qy 993 TCCTCTGAAAGATAATCCCTTACCTGTGACTGTGAGAGAACCTGAGAAAG 1052

Db 191 TCCTGGAAAGATAATCCCTTACCTGTGACTGTGAGAGAACCTGAGAAAG 250

Qy 1053 AGCTGGACAGCTCCCTCTGTTCTGAAGTGCACAGGGAGGGTCTCTGAGTG 1112

Db 251 AGTGGAAAGCTCCCTCTGTTCTGAAGTGCACAGGGAGGGTCTCTGAGTG 310

Qy 1113 AGGGTCTCGGGAGGTCCCTCAGCTGCTTCTAATGAGGCTGCTCTTGG 1172

Db 311 AGGGTCTCGGGAGGTCCCTCAGCTGCTTCTAATGAGGCTGCTCTTGG 370

Qy 1113 ATGATGGTAG 1183

Db 371 ATGATGGCTAG 381

Search completed: February 20, 2004, 01:41:31  
 Job time : 7840.77 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:32:00 ; Search time 191.966 Seconds  
(Without alignments)  
3419.919 Million cell updates/sec

Title: US-09-939-853A-74  
Perfect score: 1.83  
Sequence: 1 agcttagctcaaggacc...tctttggatgtccatg 1183

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgcn2\_6\_ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgcn2\_6\_ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgcn2\_6\_ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgcn2\_6\_ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgcn2\_6\_ptodata/2/ina/PCMS\_COMB.seq: \*  
6: /cgcn2\_6\_ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	123	10.4	2015	4	US-09-023-655-1105	Sequence 1105, AP
2	119.4	10.1	2298	4	US-09-023-655-1158	Sequence 1158, AP
3	11.0	9.3	2354	4	US-09-023-655-1080	Sequence 1080, AP
4	107.4	9.1	2129	4	US-09-023-116-1452	Sequence 1452, AP
5	101	8.5	675	1	US-09-707-793A-3	Sequence 3, Appli
6	101	8.5	675	1	US-09-707-792A-3	Sequence 3, Appli
7	92.6	7.8	2435	4	US-09-023-655-11313	Sequence 1313, AP
8	91	7.7	2647	4	US-09-220-132-77	Sequence 77, Appli
9	90.8	7.7	1611	1	US-07-820-011A-3	Sequence 77, Appli
10	90.8	7.7	1611	4	US-09-60-73-3	Sequence 3, Appli
11	71	6.0	1491	2	US-09-006-675-1	Sequence 3, Appli
12	69.6	7.6	1626	4	US-09-860-473-10	Sequence 10, Appli
13	80.4	6.8	1602	1	US-07-820-011A-1	Sequence 1, Appli
14	80.4	6.8	1602	5	PCT-US93-0045-1	Sequence 1, Appli
15	80.4	6.8	1759	4	US-09-470-881-2	Sequence 2, Appli
16	71	6.0	1491	3	US-09-028-603A-1	Sequence 1, Appli
17	68.4	5.8	282	2	US-09-096-675-5	Sequence 1, Appli
18	68.4	5.8	282	3	US-09-228-603A-5	Sequence 5, Appli
19	60.4	5.1	4517	4	US-09-470-881-7	Sequence 7, Appli
20	60.4	5.1	4517	5	PCT-US93-06251-83	Sequence 83, Appli
21	59.4	5.0	874	4	US-09-023-655-931	Sequence 931, Appli
22	55.8	4.7	1467	4	US-09-579-182-2	Sequence 2, Appli
23	55.8	4.7	1548	4	US-09-099-053-1	Sequence 1, Appli
24	46	4.1	144	5	PCT-US93-06251-13	Sequence 13, Appli
25	46	3.9	190	5	PCT-US93-06251-14	Sequence 14, Appli

#### ALIGNMENTS

RESULT 1 US-09-023-655-1105						
/ Sequence 1105, Application US/09023655						
/ Patent No. 6807879						
/ GENERAL INFORMATION:						
/ APPLICANT: Cocks, Benjamin G.						
/ APPLICANT: Susan G. Stuart						
/ APPLICANT: Jeffrey J. Seihammer						
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE						
/ NUMBER OF SEQUENCES: 1508						
/ CORRESPONDENCE ADDRESS:						
/ ADDRESSE: INCYTE PHARMACEUTICALS, INC.						
/ STREET: 3174 PORTER DRIVE						
/ CITY: PALO ALTO						
/ STATE: CALIFORNIA						
/ COUNTRY: USA						
/ ZIP: 94304						
/ COMPUTER READABLE FORM:						
/ MEDIUM TYPE: Floppy disk						
/ COMPUTER: IBM PC compatible						
/ OPERATING SYSTEM: PC DOS/MS-DOS						
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2						
/ CURRENT APPLICATION DATA:						
/ APPLICATION NUMBER: US/09/023, 655						
/ FILING DATE: HEREWITHE						
/ CLASSIFICATION:						
/ PRIOR APPLICATION DATA:						
/ APPLICATION NUMBER:						
/ PILING DATE:						
/ CLASSIFICATION:						
/ ATTORNEY/AGENT INFORMATION:						
/ NAME: Zeller, Karen J.						
/ REGISTRATION NUMBER: 37, 071						
/ REFERENCE/DOCKET NUMBER: PA-0001 US						
/ TELECOMMUNICATION INFORMATION:						
/ TELEPHONE: (650) 845-0555						
/ TELEFAX: (650) 845-1166						
/ INFORMATION FOR SEQ ID NO: 11.05:						
/ SEQUENCE CHARACTERISTICS:						
/ LENGTH: 2015 base pairs						
/ TYPE: nucleic acid						
/ STRANDEDNESS: single						
/ TOPOLOGY: linear						
/ IMMEDIATE SOURCE:						
/ LIBRARY: GENBANK						
/ CLONE: 9183911						
/ US-09-023-655-1105						
/ Query Match						
10.4%; Score 123; DB 4; Length 2015;						

Patent No. 5219739  
Sequence 28, Appli  
Sequence 17, Appli  
Sequence 70, Appli  
PCT-US93-06251-0  
PCT-US95-10973A-18  
US-09-244-883-13  
US-09-392-932-6  
US-09-574-708A-1  
US-09-392-931-1  
PCT-US95-11973A-68  
PCT-US95-10973A-86  
US-08-718-904-1  
US-09-449-249-1  
PCT-US95-10973A-25  
US-09-244-583-5  
Patent No. 514596  
Sequence 1, Appli

Best Local Similarity 53.9%; Pred. No. 1.5e-25;  
 Matches 283; Conservative 0; Mismatches 230; Indels 12; Gaps 1;

Qy CCAAGCCAGGACCTGTGACCATGAAAGAGAACGGCAGCCCTGGCGCT 516  
 Db CAACGCAAACACACAGGAATCAGGGAGGGCTCTGAGACATCATGTGTTGCCCT 362  

Qy GGGCACTTTCGGAGGTGCCCCGCGAGCTGNGCTGAGACTGGGAGCACATGAC 576  
 Db GTATATTACAGGGCATTACCGAAGACCTAACGTCAGAGGGGACAGATGGT 422  

Qy CATCGCTCTCAGGATGGAGACTGGTGGAGGAGTGTGAAGTCTAGGGAGAGATA 636  
 Db GGTCATAGGAATCGGGAGTGTGGAGGAGTCATCTGGCCACCGGAGAGGG 482  

Qy TAACATCCCAGGNCACCTGGGAAACCTCCAT-----GGGGGCTGTA 684  
 Db CTACATCCAGCAACTATCGCCGCGTGAAGGGGGTTTT 542  

Qy TGAGGCCTGAGCAGGGAGAAAGCAGAGGAAGCTGTGTTACCTGGAAACCTGGAGS 744  
 Db CAAGGECATAGCCGAAGAACGCGAGGCCAACACTGCTGCTGCCCAATGCGGG 602  

Qy GGCCTTCCTCATCCGGAGCCAGACCGAGGGCTCTACTCTCTGTCASTCCGCT 804  
 Db CCTCTTCATGATCCGGATAGCGAGACCTAAAGGAGCTACTCTTGTGCGAGA 662  

Qy CAGCCCCCTGCACTGTGGACCGATAACCTACAGATCGACATCGCTGACATGG 864  
 Db CTACGACCCCTGGCAGGGAGATAACCGTGAACATTAACAGATCGACATCGCTGACAGG 722  

Qy CTGGCTGTATACTCTACCGGGCTCACCTCCCTACTCAGCCCTGGGACCTTA 924  
 Db GGGCTTCATATATECCCCGAAGCACCTTCAAGACTCTGAGGCTGTGGACCACTA 782  

Qy CTCTGAGCTGGGATGACATCTGTGCCCTACTCAAGGGGCCCTG 969  
 Db CAAGAGGGAAACGAGGGCTCTGCGAGAACTGTGGCCCTG 827

## RESULT 2

US-09-023-655-1158

Sequence 1158 Application US/09023655

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023-655

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071  
 ; REFERENCE/DOCKET NUMBER: PA-0001 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 855-0555  
 ; FAX: (650) 845-4166  
 ; INFORMATION FOR SEQ ID NO: 1158:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2298 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GENBANK  
 ; CLONE: 9187268  
 ; US-09-023-655-1158

Query Match Score 119.4; DB 4; Length 2298;  
 Best Local Similarity 53.3%; Pred. No. 1.8e-24;  
 Matches 283; Conservative 0; Mismatches 236; Indels 12; Gaps 1;

Qy 454 TGTCCGAGGCCAGGGACCTGTGACCATGGAGACAGAGAACGGCCACAGCGGTGGC 513  
 Db 447 TTACCTGGAGAGGTTAACATAGATCCAGAGAACAGAGACATTGGTAGC 506  

Qy 514 CCTGGGCAGTTCCGGCAGETGGGCCGGCGAGCTGGGGAGCCATT 573  
 Db 507 CTGGACCCCTATGATGGATGCTGTTACAGAAAGGAGAGAT 566  

Qy 574 GACCAGTGCTCTGAGGATGAGCTGGTGAAGCTGCTGAAGCTCAGSCAGAGA 633  
 Db 567 GAAAGCTGCTGGAGGAGATGAGATGGTGAAGCAAGCTCCCTTACAAAAAGA 626  

Qy 634 GTATAACATCCCAGGTCACCTGGCAACTCTACACCTTAAAGTCTCC 681  
 Db 627 AGGCCTCATCCCGAGCAACTATGTCAGGCAACTCTACACCTTAAAGAGTGGTT 686  

Qy 682 GTATGGGGCTGATCCGAGGAAAGCAGGAGGACTCTGTAGTCGG 741  
 Db 687 TTTCARGATAAACTGGAGGAGGAGGAGCTTGGACATAGGAAATAGCGC 746  

Qy 742 AGGGGCTTCTCATCCGGAGAGGAGGAGGACTCTGTAGTCGG 801  
 Db 747 TGGAGTTCTCTTATAGAAAGTGAACATTAAGGAAGCTTCTGTCGTCAG 806  

Qy 802 CCTCAGCGCCCTGATCCGGAACTCTGAGGACTCTGTAGTCGG 861  
 Db 807 AGACTTGGCTGAGCTGAGATGCTGTTACGCTACAGGATCCACTGCTGATGTTGATA 866  

Qy 862 TGGCTGCTGACATCTACGGCTCTACCTCCCTACTCTACGGCCCTGGACCCA 921  
 Db 867 TGGGGCTTATACATCTACGGATACATCTACGGATCTACGGATATGATTAAGAACAA 926  

Qy 922 TTACGTGAGGATGACATCTGCGCCCTACTCTACGGAGACCTGT 972  
 Db 927 TTACAAAAGGAGGAGGAGCTCTGAGAAGATTGGAGAAGGCTTGTAT 977

RESULT 3  
 US-09-023-655-1080

; Sequence 1080 Application US/09023655  
 ; Patent No. 6607879  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cocks, Benjamin G.  
 ; APPLICANT: Susan G. Stuart  
 ; APPLICANT: Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: EXPRESSION  
 ; NUMBER OF SEQUENCES: 1500  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA

COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HERENWIT  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1080:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2354 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: 9182573

SUIT 4  
-09-016-434-1452  
Sequence 1452 Application US/09016434  
Petent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer

Page 4

Qy 918 ACCATTACTGTGAGCTGGGGATGACATCTGCTTACTCAAGGAGCCCTACTGTCAGGCCCTGTGTCTGC 978  
 Qy ; Sequence 3, Application US/08/07793A  
 Db ; Patent No. 5776636  
 Qy ; GENERAL INFORMATION:  
 Db ; APPLICANT: SALONE, SCOTT P.  
 Qy ; TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING  
 Db ; TITLE OF INVENTION: FUSION PROTEINS  
 Qy ; NUMBER OF SEQUENCES: 17  
 Db ; CORRESPONDENCE ADDRESS:  
 Qy ; ADDRESSEE: Merck & Co., Inc.  
 Db ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
 Qy ; CITY: Rahway  
 Db ; STATE: NJ  
 Qy ; COUNTRY: USA  
 Db ; ZIP: 07065-0900  
 Qy ; COMPUTER READABLE FORM:  
 Db ; MEDIUM TYPE: Diskette  
 Qy ; COMPUTER: IBM Compatible  
 Db ; OPERATING SYSTEM: DOS  
 Qy ; SOFTWARE: FastSEQ for Windows version 2.0  
 Db ; CURRENT APPLICATION DATA:  
 Qy ; APPLICATION NUMBER: US/08/707,793A  
 Db ; FILING DATE: 04-SEP-1996  
 Qy ; CLASSIFICATION: 435  
 Db ; PRIOR APPLICATION DATA:  
 Qy ; APPLICATION NUMBER:  
 Db ; FILING DATE:  
 Qy ; ATTORNEY/AGENT INFORMATION:  
 Db ; NAME: Camara, Valerie J.  
 Qy ; REGISTRATION NUMBER: 35,090  
 Db ; FILING NUMBER: 19494  
 Qy ; TELECOMMUNICATION INFORMATION:  
 Db ; TELEPHONE: 908-594-3902  
 Qy ; TELEFAX: 908-594-4720  
 Db ; TELEX:  
 Qy ; INFORMATION FOR SEQ ID NO: 3:  
 Db ; SEQUENCE CHARACTERISTICS:  
 Qy ; LENGTH: 675 base pairs  
 Db ; TYPE: nucleic acid  
 Qy ; STRANDEDNESS: single  
 Db ; TOPOLOGY: linear  
 Qy ; MOLECULE TYPE: Genomic DNA  
 Db ; US-08-707-793A-3

RESULT 5

Qy ; Sequence 3, Application US/08/07793A  
 Db ; Patent No. 5776636  
 Qy ; GENERAL INFORMATION:  
 Db ; APPLICANT: SALONE, SCOTT P.  
 Qy ; TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING  
 Db ; TITLE OF INVENTION: FUSION PROTEINS  
 Qy ; NUMBER OF SEQUENCES: 17  
 Db ; CORRESPONDENCE ADDRESS:  
 Qy ; ADDRESSEE: Merck & Co., Inc.  
 Db ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
 Qy ; CITY: Rahway  
 Db ; STATE: NJ  
 Qy ; COUNTRY: USA  
 Db ; ZIP: 07065-0900  
 Qy ; COMPUTER READABLE FORM:  
 Db ; MEDIUM TYPE: Diskette  
 Qy ; COMPUTER: IBM Compatible  
 Db ; OPERATING SYSTEM: DOS  
 Qy ; SOFTWARE: FastSEQ for Windows version 2.0  
 Db ; CURRENT APPLICATION DATA:  
 Qy ; APPLICATION NUMBER: US/08/707,793A  
 Db ; FILING DATE: 04-SEP-1996  
 Qy ; CLASSIFICATION: 435  
 Db ; PRIOR APPLICATION DATA:  
 Qy ; APPLICATION NUMBER:  
 Db ; FILING DATE:  
 Qy ; ATTORNEY/AGENT INFORMATION:  
 Db ; NAME: Camara, Valerie J.  
 Qy ; REGISTRATION NUMBER: 35,090  
 Db ; FILING NUMBER: 19494  
 Qy ; TELECOMMUNICATION INFORMATION:  
 Db ; TELEPHONE: 908-594-3902  
 Qy ; TELEFAX: 908-594-4720  
 Db ; TELEX:  
 Qy ; INFORMATION FOR SEQ ID NO: 3:  
 Db ; SEQUENCE CHARACTERISTICS:  
 Qy ; LENGTH: 675 base pairs  
 Db ; TYPE: nucleic acid  
 Qy ; STRANDEDNESS: single  
 Db ; TOPOLOGY: linear  
 Qy ; MOLECULE TYPE: Genomic DNA  
 Db ; US-08-707-793A-3

Query Match 8.5% Score 101; DB 1; Length 675;  
 Best Local Similarity 59.0%; Pred. No. 2e-19;  
 Matches 173; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 677 TGGCTGTATGAGGCCCTGAGCAGGAGAAACAGAGGAACTGCTGTGTTACCTCGGAACT 935  
 Db 373 TGGTCTCAAGAACCTGAGCCGAAGAGGAGGGCTCTACTCTGTCTGGCGGGAC 432  
 Qy 737 CCNGAGGGGCCCTCCATCCGGAGGAGCCAGAACCTGAGGGCTCTACTCTGTCTGGCGGGAC 796  
 Db 433 ACCTACGGTCTCCATCCGGAGAACCTGAGGGCTCTACTCTGTCTGGCGGGAC 492  
 Qy 797 GTCCGCCCTCAAGGCCCTGGTGAATCTCACCCGGCTTACCTGGACCCGATGAGACTACGGATCCACTGCCCT 856  
 Db 493 GTCCGGGAACCTTCACCAAGAACCTGAGGGAGGGCTCTACTCTGTCTGGCGGGAC 552  
 Qy 857 GACATGATGGTGGTGTGATCTCACCCGGCTTACCTGGACCCGATGAGACTACGGATCCACTGCCCT 916  
 Db 553 GACATGATGGTGGTGTGATCTCACCAAGAACCTGAGGGAGGGCTCTACTCTGTCTGGCGGGAC 612

```

RESULT 6
US-08-707-792A-3
Sequence 3, Application US/08707792A
Patent No. 5783398
GENERAL INFORMATION:
APPLICANT: MARCY, ALICE
APPLICANT: SALOME, SCOTT P.
APPLICANT: WISNIERSKI, DOUGLAS
TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-9900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707, 792A
FILING DATE: 04-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Camare, Valerie J
REGISTRATION NUMBER: 35, 090
REFERENCE/DOCKET NUMBER: 19524
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3902
TELEFAX: 908-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 675 base Pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-707-792A-3

Query 917 GACCATTACTGAGCTGGGGATGACATCTGCCTACTCAAGGGCCCTG 969
Db 613 CGCCATTAACAAATGCTAGATGGGTGACACGGTGAACGGCCCTG 665

Query 917 GACATCGGTCCTCTCATCGGAGAACGGAGCTCTACTCTGTG 857
Db 433 ACTACGGTCCTCTCATCGGAGAACGGAGCTCTACTCTGGCTG 857
Query 797 GTCCGGCTCTGGCGCTGATCTCTGGACGGATCAGGACTAACGGTCCACTCGCTG 553
Db 493 GTCCGGGACTTGACGACCAACGGAGGGAGGTGCTGAACTAACGGTCAATCGCTG 553
Query 737 CCTGGAGGGCCCTCTCATCGGAGAACGGAGCTCTACTCTGTG 857
Db 433 ACTACGGTCCTCTCATCGGAGAACGGAGCTCTACTCTGGCTG 857
Query 737 TGTTCTTCARGAACCTGACGCCAGAACCGGGAGGTGCTGAACTAACGGTCAATCGCTG 553
Db 493 GTCCGGGACTTGACGACCAACGGAGGGAGGTGCTGAACTAACGGTCAATCGCTG 553
Query 737 GACAATGGCTGGCTGACCTCACTGGGAGAACGGAGCTCTACTCTGTG 857
Db 433 ACTACGGTCCTCTCATCGGAGAACGGAGCTCTACTCTGGCTG 857
Query 737 TGTTCTTCARGAACCTGACGCCAGAACCGGGAGGTGCTGAACTAACGGTCAATCGCTG 553
Db 493 GTCCGGGACTTGACGACCAACGGAGGGAGGTGCTGAACTAACGGTCAATCGCTG 553

```

Qy 917 GACCATTACTCTAGCTGGGATGACATCTGCTACTCAAGGAGCCCTG 969  
 Db 613 CGCCATTACCAATGCTTCATGGCTCACAGGTGAGGCCCTG 665

RESULT 7  
 US-09-023-655-1313  
 Sequence 1.3.13, Application US/09023655  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cocks, Benjamin G.  
 ; PATENT NO.: 6506507  
 ; PRIORITY NUMBER: US 09/220,132  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
 ; NUMBER OF SEQUENCES: 1508  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/023, 655  
 ; FILING DATE: HEREMITH  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeller, Karen J.  
 ; REGISTRATION NUMBER: 37,071  
 ; REFERENCE DOCKET NUMBER: PA-0001 US  
 ; TELEPHONE: (650) 855-0555  
 ; TELEFAX: (650) 855-4166  
 ; INFORMATION FOR SEQ ID NO: 1313:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2435 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GENBANK  
 ; CLONE: 9338227  
 ; US-09-023-655-1313

Qy 912 TGGTGGACATTACTCTAGCTGGGATGACATCTGCTACTCAAGGAGCCCTG 970  
 Db 1050 TTGTACACATATTACTAGAGCTGAGCTGGCTTAGTAGTTCCTGT 1108

RESULT 8  
 US-09-220-132-77  
 ; Sequence 7.7, Application US/09220132  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shyjin, Andrew W.  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER  
 ; FILE REFERENCE: 07334-074001  
 ; CURRENT APPLICATION NUMBER: US/09/220,132  
 ; CURRENT FILING DATE: 1998-12-23  
 ; PRIOR APPLICATION NUMBER: US 60/079,303  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: US 60/068,821  
 ; PRIOR FILING DATE: 1997-12-24  
 ; NUMBER OF SEQ ID NOS: 191  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO: 77  
 ; LENGTH: 2647  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-220-132-77

Qy Query Match 7.7%; Score 91; DB 4; Length 2647;  
 Db Best Local Similarity 56.5%; Pred. No. 2.9e-16;  
 Matches 169; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy 672 ATGGGTGGCTGTATGAGGGCTGAGGAGAAAGGAGAAAGGAGAAACTCGTGTGTACCTG 731  
 Db 1019 AAGAGTGTTACTTTGAAACTTGGCGAAAGATGTTGGGACAGCTATTGTCCTTG 1078  
 Qy 732 GGACCCCTGGAGGGCTCTCCATCGGGAGAGCCAGAACGGGCCTACTCTC 791  
 Db 1079 GAAACCCAGAGGTACCTCTTATCGGAGTAAACCAAGGAGTAAACCAAGGCTATTCAC 1138  
 Qy 792 TGTCAAGTCGGCTCAAGGGATCAGACACTACAGGATCCACT 851  
 Db 1139 TTCTCATCGTGAATGGGTATGATGAAAGGACCATGTTATAAATTCGCA 1198  
 Qy 852 GCCTTGACAAATGGCTGTATCATCTACGGGCTCATCCCTCATCCAGCCC 911  
 Db 1199 AACTTGACAAATGGTATCATACATTCACCCAGTTGAAACACTTCAGCAGC 1258  
 Qy 912 TGTGGACCAATTAATGCTGAGCTGGCGATGACATCGCTGCTACTCAAGGAGCCCTG 970  
 Db 1259 TGTACACATTAATGAGAGCTGAGCTGCTCTGCGCTTAGTAGTTCCTGT 1317

RESULT 9  
 PCT-US93-06251-77  
 ; Sequence 7.7, Application PC/US9306251  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wickstrom, Eric and Rife, Jason P.  
 ; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing Stereospecific Alkylphosphonates and Arylphosphonates  
 ; NUMBER OF SEQUENCES: 93  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCULLY, MURPHY & PRESSER  
 ; STREET: 400 Garden City Plaza  
 ; CITY: Garden City  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 11530  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US93/06251  
 FILING DATE: 19930630  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: DIGISIO, Frank S.  
 REGISTRATION NUMBER: 31,346  
 REFERENCE/DOCKET NUMBER: 8586  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 516-742-4343  
 TELEX: 230 901 SANS UR  
 INFORMATION FOR SEQ ID NO: 77:  
 LENGTH: 2647 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: Linear  
 MOLECULE TYPE: DNA (genomic)  
 PCT-US93-06251-77

Query Match 7.5%; Score 91; DB 5; Length 2647;  
 Best Local Similarity 56.5%; Pred. No. 2.9e-16;  
 Matches 169; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy	672	ATGGGTGGCTGTATAGGGCCTGANGCAGGAAAGCAGAGAAACTCTGTGTGTTACCTG	731
Db	1019	AAGAGTGGTAACTTGGAAACATTGGCTGAAAGATGAGCTGGTGAGGACAGCTTGTCCCTTG	1078
Qy	732	GGAAGCCTGGAGGGCCCTTCATCCGGAGGCCAACACAGGAGGCTCTTACTCTC	791
Db	1079	GAAACCCAAAGGTAACTCTTCTTATCCGGAGACTAAACACCAAAAGGTGCTATCAC	1138
Qy	792	TGTCACTGCCCTTAAGCCGGCTGATCTGGGACCGATCACACACTACAGATCACT	851
Db	1139	TTTCTATCCCTGATGGGATGATATGAAAGGAAACCATTAATTAATTCGCA	1198
Qy	852	GCCPTGACAATGGCTGGCTGACATCTCACGCCCTCACCTCCCTCACCTCCAGGCC	911
Db	1199	AACPTGACAATGGCTGGCTGACATCTCACGCCCTCACCTCCCTCACCTCCAGGCC	1258
Qy	912	TGGTGACCATTACTCTGAGCTGGGGATGACATCTGCTGCTACTAAGGAGCCCTGT	970
Db	1259	TGTGACAACTACTAGAGAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1317

RESULT 10  
 / Sequence 3, Application US/07820011A  
 / Patent No. 5336615  
 / GENERAL INFORMATION:  
 / APPLICANT: Bell, Leonard  
 / APPLICANT: Madri, Joseph A.  
 / APPLICANT: Warren, Stephen L.  
 / APPLICANT: Luthringer, Daniel J.  
 / TITLE OF INVENTION: Genetically Engineered  
 / Endothelial Cells Exhibiting Enhanced  
 / Title of Invention: Migration  
 / Title of Invention: and Plasminogen Activator Activity  
 / Number of Sequences: 4  
 / Correspondence Address:  
 / Addressee: Maurice M. Klee  
 / Street: 1951 Burr Street  
 / City: Fairfield  
 / State: Connecticut  
 / Country: USA  
 / Zip: 06430  
 / Computer Readable Form:  
 / Medium Type: 5.25 inch, 360 Kb storage  
 / Computer: IBM PC XT  
 / Operating System: PC-DOS/MS-DOS 2.10  
 / Software: Displaywrite 3

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/820,011A  
 FILING DATE: 19930106  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Klee, Maurice M.  
 REGISTRATION NUMBER: 30,399  
 REFERENCE/DOCKET NUMBER: LB-101  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (103) 255 14 00  
 TELEFAX: (203) 254 1101  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1611  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: Double  
 TOPOLOGY: Linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: No  
 ANTI-SENSE: No  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapien  
 POSITION IN GENOME:  
 CHROMOSOME SEGMENT: Chromosome 20  
 PUBLICATION INFORMATION:  
 AUTHORS: Anderson, Stephen K.  
 AUTHORS: Gibbs, Carol P.  
 AUTHORS: Tanaka, Akio  
 AUTHORS: Kung, Hsing-Jien  
 AUTHORS: Fujita, Donald J.  
 TITLE: Human Cellular src Gene:  
 TITLE: Nucleotide Sequence and Derived Amino Acid Sequence of the Region Coding for the Carboxy-Terminal Two-Thirds of the Protein  
 TITLE: pp60-src  
 JOURNAL: Molecular and Cellular Biology  
 VOLUME: 5  
 ISSUE: 5  
 PAGES: 1122-1129  
 DATE: May, 1985  
 PUBLICATION INFORMATION:  
 AUTHORS: Tanaka, Akio  
 AUTHORS: Gibbs, Carol P.  
 AUTHORS: Arthur, Richard R.  
 AUTHORS: Anderson, Stephen K.  
 AUTHORS: Kung, Hsing-Jien  
 AUTHORS: Fujita, Donald J.  
 TITLE: DNA Sequence Encoding the Amino-Terminal Region of the Human c-src  
 TITLE: Protein: Implications of Sequence Divergence among src-Type Kinase  
 TITLE: Oncogenes  
 JOURNAL: Molecular and Cellular Biology  
 VOLUME: 7  
 ISSUE: 5  
 PAGES: 1978-1983  
 DATE: May, 1987  
 US-07-820-011A-3

Query Match 7.7%; Score 90.8; DB 1; Length 1611;  
 Best Local Similarity 53.4%; Pred. No. 2.6e-16;  
 Matches 222; Conservative 0; Mismatches 182; Indels 12; Gaps 1;

Qy	556	GAGACTCGGGAGCCATGACCATGCTCTGAGGATGAGACTGGGAGCTGGCTGCTGCCTGCAC	615
Db	318	GAAAGCGAGGGCTCAGATGTCAAACACAGGAGACTGGCTGCTGCCTGCAC	377
Qy	616	TGAGCTCTAGGCGAGAGTATAACATCCCGAGCTCAGCTG-----GCAA	663
Db	378	GCTCGCAGGAGACAGGCTACATCCCGCAACTAGGCGCCCTGGACTCCAT	437
Qy	664	AGTCCTCGATGGGGGGGTATAAGGGCTGACCAAGGAAAGCAGGAAAGC	723

RESULT 12  
PCT-US93-00445-3  
Sequence 3, Application PC/TU9300445  
GENERAL INFORMATION:  
APPLICANT: Bell, Leonard  
APPLICANT: Madr, Joseph A.  
APPLICANT: Warren, Stephen L.  
APPLICANT: Luthinger, Daniel J.  
TITLE OF INVENTION: Genetically Engineered  
Title of Invention: Endothelial Cells  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 760 Kb storage  
COMPUTER: DELTA 486/50  
OPERATING SYSTEM: DOS 5.0  
SOFTWARE: Displaywrite 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00445  
FILING DATE: 19930105  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/820,011  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: AUX-101PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255 1400  
TELEFAX: (203) 254 1101  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1611  
TYPE: NUCLEIC ACID  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULAR TYPE: cDNA to mRNA  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE: Homo sapien  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: Chromosome 20  
PUBLICATION INFORMATION:  
AUTHORS: Anderson, Stephen K.  
AUTHORS: Gibbs, Carol P.  
AUTHORS: Tanaka, Akio  
AUTHORS: Kung, Hsing-Jien  
AUTHORS: Fujita, Donald J.  
TITLE: Human Cellular src Gene:  
TITLE: Nucleotide Sequence and Derived Amino  
Acid Sequence of the Region Coding for  
the Carboxy-Terminal Two-Thirds of  
JOURNAL: Molecular and Cellular Biology  
VOLUME: 5  
ISSUE: 5  
PAGES: 1122-1129  
DATE: MAY, 1985  
PUBLICATION INFORMATION:  
AUTHORS: Tanaka, Akio  
AUTHORS: Gibbs, Carol P.  
AUTHORS: Arthur, Richard R.  
AUTHORS: Anderson, Stephen K.  
AUTHORS: Kung, Hsing-Jien  
AUTHORS: Fujita, Donald J.

RESULT 11  
US-09-860-473-3  
Sequence 3, Application US/09860473  
Patent No. 656732  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-C EXPRESSION  
FILE REFERENCE: RTS-022  
CURRENT APPLICATION NUMBER: US/09/860,473  
CURRENT FILING DATE: 2001-05-18  
SEQ ID NO: 3  
LENGTH: 1611  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1611)  
US-09-860-473-3

Query Match Score 7.7%  
Best Local Similarity 53.4%; Pred. No. 2.6e-16;  
Matches 222; Conservative 0; Mismatches 182; Indels 12; Gaps 1;

QY 556 GAGACTCGGGGAGCCATTGACATCGTCTCTGAGATGAGACTGGTGACCGTGTGTC 615  
Db 318 GAAAGGCGAGGGCTCCAGATGTCAACACAGAACAGGGCT 377  
QY 616 TGAAGTCTAGCCGAGAGATAACATCCCAGGTCAACACAGAACAGGGCT 63  
Db 378 GCTCAGCACAGCACAGGGCTACATCCCAGCAACTACAGGGCCCTCCGACTCCAT 437  
QY 664 AGTCTCCATGGGCTTATAGGGCCCTGAGGAGGAAGGAACTGCTGTT 723  
Db 438 CAGGCTGAGAGCTGATTTGGCAAGATCACAGAGGGTCAAGGGTTACTGCT 497  
QY 724 TTAACTCTGGAAACCTGGGGCTCTCATCGGGAGACGCCAGGAGGGCTC 783  
Db 498 CAATCGAGGAACCGAGGAGCTCTGTAACGGCAAGGGCTAACGTGAGCACTACAA 617  
QY 844 GATCCACTGCCTTGACAATGGCTGGCTGAGACTCTCACCGCCCTCACCTCCCTCACT 903  
Db 618 GATCGCAAGCTGAGGGCTACCTCCACCCAGTCAAGCT 677  
QY 904 CAGGGCCCTGGGACCAATTACTCTGAGCTGGGATGACATCTGCTGCTACTCA 959  
Db 678 GCACCACTGGCTCCACTCTCACACCCGATGGCTGCAACCGCTCA 733

TITLE: DNA Sequence Encoding the  
 / Amino-Terminal Region of the Human c-src  
 / Protein: Implications of Sequence  
 / Divergence among src-Type Kinase  
 / Oncogenes  
 / JOURNAL: Molecular and Cellular Biology  
 / VOLUME: 7  
 / ISSUE: 5  
 / PAGES: 1978-1983  
 / DATE: May, 1987  
 / PCT-US93-00445-3

Query Match 7.7%; Score 90.8; DB 5; Length 1611;  
 Best Local Similarity 53.4%; Pred. No. 2.6e-16;  
 Matches 222; Conservative 0; Mismatches 182; Indels 1; Gaps 1;

Qy 556 GAGACTGGGAGCCATTGACCATGCTTGAGTGAAGCTGGCACTGGCCTAC 615  
 Db 318 GAAGGGAGGGTTCAGTTGACACACAGGGAACTGTGCGGCCACTC 377

RESULT 14

Qy 616 TGAAGTCTCAGGAGAGATAAACATCCCAGCTTCAACTGG-----GCAA 663  
 Db 378 GCTCGCAGGGACAGACAGCTTAATCCCAGAACTAGTGCGCCCTCCACTCAT 437

Qy 664 AGTCCTCCATTGGTGCCTGATGGGGCTTGGAGGGAACTGGCTGTT 723  
 Db 438 CCAGGTGAGAGTGTATTTGGAAAGATCACAGGGAGTCAGGGGTACTCT 497

Qy 724 GTTAATCTGGAAACCTGGAGGGCTTCATACGGAGAGCAGGAGGCT 783  
 Db 498 CAATGGAGAAACCGAGGGACCTTCTGGAGAAGTGAACCAAGTGC 557

Qy 784 TTACTCTCTGTCAGTCGGCTCACCGGCTGATCCGGAACTACAG 843  
 Db 558 CTACTGCCPCTCAAGTGACTCTGATAACGGCAAGCTGAACNTACA 617

Qy 844 GATCCACTGGTGGACCATACTCTGAGCTGGCGATGACATOTCAACGCTTCCCTCACT 903  
 Db 618 GATCCGCAAGCTGGAGGGGCTTCATACACCTCCGGACCCAGTAAACGCT 677

Qy 904 CCAGGCCCTGGTGGACCATACTCTGAGCTGGCGATGACATCTGCTGCCACTCA 959  
 Db 678 GCAGCAGCTGGTGGCTTACTCAAAACCGCGATEGGCTGTGCAACGGCTCA 733

RESULT 13

US-09-860-473-10  
 / Sequence 10. Application US/09860473  
 / Patent No. 6656732

/ GENERAL INFORMATION:  
 / APPLICANT: C. Frank Bennett  
 / TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-C EXPRESSION  
 / FILE REFERENCE: RIS-0222  
 / CURRENT APPLICATION NUMBER: US/09/860,473  
 / CURRENT FILING DATE: 2001-05-18  
 / NUMBER OF SEQ ID NOS: 169  
 / SEQ ID NO 10  
 / LENGTH: 1626  
 / TYPE: DNA  
 / ORGANISM: Mus musculus  
 / FEATURE: CDS  
 / NAME/KEY: CDS  
 / LOCATION: (1) . . . (1626)

Query Match 7.6%; Score 89.6; DB 4; Length 1626;  
 Best Local Similarity 56.9%; Pred. No. 5.8e-16;  
 Matches 164; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 672 ATGGTGGTGTATGGGGCTAGCAAGGAGAACTGCTGTGTACCTG 731  
 Db 461 AGCAAGTGGTACCTTGGCAAGGATCACTAGACGGGAATAGAECGGTGTCAACGCCG 520

Qy 732 GGAAACCTGGAGGGCCTTCCATCGGGAGAGCGACAGGAGGGCTTACTCTC 791  
 Db 521 AGAACCGAGGGACCTTCTGTAAGGAGTAAAGGCTACTGCC 580

Qy 792 TGTCACTCCGCCTCAGCGCCTGATCTGGACGGATCAAGACATAGGGTCACT 851  
 Db 581 TCTCTGTATCGACTGACATGGCAAGCCCTAAATGTAAAACACTACAGATCCGCA 640

Qy 852 GCCTGACATGGCTGGCTGAGCTGACATCTACGGGCTCACCTCCCTCACTCCAGGCC 911  
 Db 641 AGCTGGACAGGGCGSTTTCATACACATGTTCAACAGCTGGAGCAG 700

Qy 912 TGTGAGGACATTACTTGAATCTGAGTGGGATGACATCTGTSCTACTCA 959  
 Db 701 TGCTGGCTTACTACCCAAATGTGATGGCTGTGACCCGCTCA 748

/ GENERAL INFORMATION:  
 / APPLICANT: Bell, Leonard  
 / APPLICANT: Maier, Joseph A.  
 / APPLICANT: Warren, Stephen L.  
 / APPLICANT: Lathriinger, Daniel J.  
 / TITLE OF INVENTION: Genetically Engineered  
 / Endothelial Cells Exhibiting Enhanced  
 / Migration  
 / Title of Invention: Migration  
 / Title of Invention: and Plasminogen Activator Activity  
 / Number of Sequences: 4  
 / Correspondence Address:  
 / Addressee: Maurice M. Klee  
 / Street: 1951 Burr Street  
 / City: Fairfield  
 / State: Connecticut  
 / Country: USA  
 / Zip: 06430  
 / Computer Readable Form:  
 / Medium Type: 5.25 inch, 360 Kb storage  
 / Computer: IBM PC XT  
 / Operating System: PC-DOS/MS-DOS 2.10  
 / Software: Displaywrite 3  
 / Current Application Data:  
 / Application Number: US/07/820,011A  
 / Filing Date: 19920106  
 / Classification: 435  
 / Attorney/Agent Information:  
 / Name: Klee, Maurice M.  
 / Registration Number: 30,399  
 / Reference/Doctet Number: LB-101  
 / Telecommunication Information:  
 / Telephone: (203) 255 1400  
 / Telefax: (203) 254 1101  
 / Information for Seq ID No: 1:  
 / Sequence Characteristics:  
 / Length: 1602 base pairs  
 / Type: NUCLIC ACID  
 / Strandedness: Double  
 / Topology: Linear  
 / Molecule Type: cDNA to mRNA  
 / Hypothetical: No  
 / Anti-Sense: No  
 / Original Source:  
 / Organism: Gallus, gallus  
 / Publication Information:  
 / Authors: Takeya, Tatsuo  
 / Authors: Hanafusa, Hideyaburo  
 / Title: Structure and Sequence of the  
 / Title: Cellular Gene Homologous to the RSV src  
 / Title: Gene and the Mechanism for Generating the  
 / Title: Transforming Virus

JOURNAL: Cell  
 VOLUME: 32  
 PAGES: 881-890  
 DATE: March, 1983  
 US-07-820-011A-1

Query Match      Score 6.8%;      Score 80.4;      DB 1;      Length 1602;  
 Best Local Similarity 54.4%;      Pred. No. 2.6e-13;  
 Matches 162;      Conservative 0;      Mismatches 136;      Indels 0;      Gaps 0;

Qy      672 ATGGGTGGCTGTATGAGGGAAAGGAGAAACTGCTGTTAACCTG 731  
 Db      437 AAGGTGGTACATTGGAGATACTCTTGCGAGTCGAGGCGCTGCCTCAACCCG 496  
 Qy      732 GGAAACCTGGAGGGCCTCTCATCCGGAGGCCAACAGGAGCTTACTCTC 791  
 Db      497 AAAACCCGGGGAACTCTCTGGCCGAGGAGCACAAAGGGCTATGCC 556  
 Qy      792 TGTAGTCGGCTCAAGCGGATCAAGACAATGTGAAGCACTACAGATCGCA 851  
 Db      557 TCCTCGGTTCTGACTTTGACAAGCCAAAGGGCTCAATGTGAAGCACTACAGATCGCA 616  
 Qy      852 GCTTGTGAACTATGGTGGCTGATCTGGCTCCCTCACTCCAGGCC 911  
 Db      617 AGCTGGACAGGGGGCTCTACATCCCTACGGACAGTCAGCTGAGCAGC 676  
 Qy      912 TGTTGGACATTAATCTGAGCTGGGGATGACATCTGCTGCCCTACTCAAGGAGCCCTG 969  
 Db      677 TGTTGGCTTACTACTCAAAATGCTGATGGTGTGCCAAAGGCTGACCAACGTTGTG 734

RESULT 15  
 PCT-US93-00445-1  
 ; Sequence 1, Application PC/TUS9300445  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell, Leonard  
 ; APPLICANT: Madri, Joseph A.  
 ; APPLICANT: Warren, Stephen L.  
 ; APPLICANT: Luthninger, Daniel J.  
 ; TITLE OF INVENTION: Genetically Engineered  
 ; NUMBER OF SEQUENCES: 4  
 ; ADDRESSEE: Maurice M. Klee  
 ; STREET: 1951 Burr Street  
 ; CITY: Fairfield  
 ; STATE: Connecticut  
 ; COUNTRY: USA  
 ; ZIP: 06430

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 760 Kb storage  
 COMPUTER: DELL 486/150  
 OPERATING SYSTEM: DOS 5.0  
 SOFTWARE: Displaywrite 3  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US93/00445  
 FILING DATE: 19930105  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/820, 011  
 FILING DATE: 06-JAN-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Klee, Maurice M.  
 REGISTRATION NUMBER: 30,399  
 REFERENCE/DOCKET NUMBER: ALX-101-PCT  
 TELECOMMUNICATION INFORMATION  
 TELEPHONE: (203) 255 1400  
 TELEX/FAX: (203) 254 1101  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1602 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: Double

Result No.	Score	Query Match	Length	DB ID	Description
1	784.4	66.3	786	14 US-10-043-649-1	Sequence 1, Appli
2	775.4	64.5	864	10 US-09-814-353-21302	Sequence 21302, A
3	758.2	64.1	763	9 US-09-867-550-953	Sequence 953, APP
4	348	29.4	444	9 US-09-867-550-951	Sequence 951, APP
5	341	28.8	875	9 US-09-867-550-915	Sequence 1915, APP
6	213.4	18.0	320	10 US-09-814-353-17314	Sequence 17314, APP
7	157.4	13.3	2665	9 US-09-954-456-499	Sequence 499, APP
8	157.4	13.3	3756	13 US-10-002-6-91	Sequence 91, APP
9	141.8	12.0	432	9 US-09-864-761-2829	Sequence 2829, APP
10	141.8	12.0	448	9 US-09-864-761-15513	Sequence 15513, APP
11	131.6	11.1	152	10 US-09-814-353-4631	Sequence 4631, APP
12	131.6	11.1	152	10 US-09-814-353-10930	Sequence 10930, APP
13	124.6	10.5	2343	15 US-10-063-674-2038	Sequence 2038, APP
14	123	10.4	1924	15 US-10-129-720-1	Sequence 1, APP
15	123	10.4	2015	9 US-09-954-456-1983	Sequence 1983, APP

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	784.4	66.3	786	14 US-10-043-649-1	Sequence 1, Appli
2	775.4	64.5	864	10 US-09-814-353-21302	Sequence 21302, A
3	758.2	64.1	763	9 US-09-867-550-953	Sequence 953, APP
4	348	29.4	444	9 US-09-867-550-951	Sequence 951, APP
5	341	28.8	875	9 US-09-867-550-915	Sequence 1915, APP
6	213.4	18.0	320	10 US-09-814-353-17314	Sequence 17314, APP
7	157.4	13.3	2665	9 US-09-954-456-499	Sequence 499, APP
8	157.4	13.3	3756	13 US-10-002-6-91	Sequence 91, APP
9	141.8	12.0	432	9 US-09-864-761-2829	Sequence 2829, APP
10	141.8	12.0	448	9 US-09-864-761-15513	Sequence 15513, APP
11	131.6	11.1	152	10 US-09-814-353-4631	Sequence 4631, APP
12	131.6	11.1	152	10 US-09-814-353-10930	Sequence 10930, APP
13	124.6	10.5	2343	15 US-10-063-674-2038	Sequence 2038, APP
14	123	10.4	1924	15 US-10-129-720-1	Sequence 1, APP
15	123	10.4	2015	9 US-09-954-456-1983	Sequence 1983, APP

\* Query Match  
 Result No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## ALIGNMENTS

```
RESULT 1
US-10-043-649-1
; Publication No. US2003059924A1
; GENERAL INFORMATION:
;   APPLICANT: Holland, Sacha J.
;   APPLICANT: Mendenhall, Marcy K.
;   APPLICANT: Pardo, Jorge
;   APPLICANT: Spencer, Collin
;   APPLICANT: Fu, C. Alan
;   APPLICANT: Luo, Ying
;   APPLICANT: Payan, Donald G.
;   APPLICANT: Mancebo, Helena S.Y.
;   APPLICANT: Wu, Jun
;   APPLICANT: Zhou, Xulan
;   APPLICANT: Shen, Mary
;   APPLICANT: Liao, X. Charlene
;   APPLICANT: Sheong, Ning
; TITLE OF INVENTION: Cloning of a No. US2003059924A1 Inhibitor of Antigen-receptor :
; FILE REFERENCE: A-70219-1.RMS/DHR
; CURRENT APPLICATION NUMBER: US-10-043-649
; CURRENT FILING DATE: 2002-01-10
; PRIORITY FILING DATE: 2001-01-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 1
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: (1)..(786)
; OTHER INFORMATION:
; US-10-043-649-1
; Query Match
```





Db 301 GCGAGCTGCTGAGACTGGGAGCCATTGACCATCCTCTGTGAG 348  
 RESULT 5  
 US-09-867-550-1915  
 Sequence 1915, Application US/09867550  
 Patent No. US20020082206A1  
 GENERAL INFORMATION:  
 / APPLICANT: Leach, Martin D.  
 / APPLICANT: Mehraban, Fuad.  
 / APPLICANT: Conley, Pamela  
 / APPLICANT: Law, Debbie  
 / APPLICANT: Topper, James  
 / TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and  
 / FILE REFERENCE: 21402-013 (Curia-313)  
 / CURRENT APPLICATION NUMBER: US/09/867,550  
 / FILING DATE: 2001-09-20  
 / PRIOR APPLICATION NUMBER: USSN 60/208,427  
 / PRIOR FILING DATE: 2000-05-30  
 / NUMBER OF SEQ ID NOS: 2125  
 / SOFTWARE: FastSEQ for Windows Version 4.0  
 / SEQ ID NO 1915  
 / LENGTH: 875  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / NAME/KEY: misc\_feature  
 / LOCATION: (1)  
 / OTHER INFORMATION: Wherein n is one of a or t or c or g  
 / US-09-867-550-1915  
 Query Match 28.8%; Score 341; DB 9; Length 875;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-94;  
 Matches 341; Conservative 0; Wimatches 0; Indels 0; Gaps 0;  
 Qy 843 GGATCCACTCCCTTGACATGGCTGCTACATCTCACCCGCCCTCAC 902  
 Db 2 GGATCCACTCCCTTGACATGGCTGCTACATCTCACCCGCCCTCAC 61  
 Qy 903 TCCAGGCCCTGGGACCATTAATCTGAGCTGGGATGACAATCTGCTGCACTGAGG 962  
 Db 62 TCCAGGCCCTGGTGGACCATTAATCTGAGCTGGGATGACAATCTGCTGCACTGAGG 121  
 Qy 963 AGCCCTGTCCTGCAAGGGCTGCCGAGATAACCCCTACCTGTGA 1022  
 Db 122 AGCCCTGTCCTGCAAGGGCTGCCGAGATAACCCCTACCTGTGA 181  
 Qy 1023 CTGTCAGGAGAACCTCAACTGGAAAGAGTGTGAAAGGCCCTCCCTGGAAAGATAACCCCTACCTGTGA 1082  
 Db 182 CTGTCAGGAGAACCTCAACTGGAAAGAGTGTGCAAGGCCCTCCCTGTGAAG 241  
 Qy 1083 CTGCACACAGGGAGCTCTTCACTGAGGTCTCTGAGATGCTGTCA 1142  
 Db 242 CTGCCACGGGGAGGAGCTCTCACTGAGGTCTGAGATGCTGTCA 301  
 Qy 1143 TCAGGCTGTAATGAGCAGGGCTGCTCTGAGATGCTGTCA 1183  
 Db 302 TCAGGCTGTAATGAGCAGGGCTGCTCTGAGATGCTGTCA 342  
 RESULT 6  
 US-09-814-353-17314  
 / Sequence 17314, Application US/09814353  
 / PUBLICATION NO. US20030165831A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Lee, John  
 / APPLICANT: Thompson, Pamela  
 / APPLICANT: Lillie, James  
 / TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 / TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 / TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
 / FILE REFERENCE: MRI-0063  
 / PRIOR APPLICATION NUMBER: US/60/235, 863



CURRENT FILING DATE: 2001-05-23  
 / PRIOR APPLICATION NUMBER: US 60/180,312  
 / PRIOR FILING DATE: 2000-02-04  
 / PRIOR APPLICATION NUMBER: US 60/207,456  
 / PRIOR FILING DATE: 2000-05-26  
 / PRIOR APPLICATION NUMBER: US 09/632,366  
 / PRIOR APPLICATION NUMBER: GB 2263.6  
 / PRIOR FILING DATE: 2000-10-04  
 / PRIOR APPLICATION NUMBER: US 60/236,359  
 / PRIOR FILING DATE: 2000-09-27  
 / PRIOR APPLICATION NUMBER: PCT/US01/00666  
 / PRIOR APPLICATION NUMBER: PCT/US01/00667  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00664  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00669  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00665  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00668  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00663  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00662  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00661  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR FILING DATE: 2000-09-21  
 / PRIOR FILING DATE: 2000-09-21  
 / PRIOR FILING DATE: 2000-06-30  
 / PRIOR FILING DATE: 2000-06-30  
 / PRIOR FILING DATE: 2001-01-29  
 / NUMBER OF SEQ ID NOS: 49117  
 / SOFTWARE: Aminoxam Sequence Listing Engine vers. 1.1  
 / SEQ ID NO: 2812  
 / LENGTH: 432  
 / TYPE: DNA  
 / FEATURE: ORGANISM: Homo sapiens  
 OTHER INFORMATION: MAP TO AL031662.24  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1  
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.4  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1  
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9  
 OTHER INFORMATION: EXPRESSED IN EPIETAL LIVER, SIGNAL = 2  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3  
 US-09-864-761-2829

Query Match 12.0%; Score 141.8; DB 9; Length 432;  
 Best Local Similarity 95.4%; Pred. No. 4.7e-33;  
 Matches 146; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 912 TGGTGGACATTACTCTGAGCTGGCGATGACATCTGCTTACTAAGAGCCGTG 971  
 Db 253 TGGGGTCCTCCTCAAGCTGGCGATGACATCTGCTTACTAAGAGCCGTG 312

QY 972 TCCTGAGAGGGCTGGCGCTTCTGCGAGATAACCCCTACCTGTGACTGTGCA 1031  
 Db 313 TCCTGAGAGGGCTGGCGCTTCTGCGAGATAACCCCTACCTGTGACTGTGCA 372

QY 1032 GGACACCACTCACTGGAAAGGTGACAGCT 1064  
 Db 373 GGACACCACTCAATGGAAAGGTGACAGCT 405

QY 972 TCCTGAGAGGGCTGGCGCTTCTGCGAGATAACCCCTACCTGTGACTGTGCA 1031

RESULT 10  
 US-09-864-761-15513  
 / Sequence 15513, Application US/09864761  
 / Patent No. US0020048763A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Pann, Sharron G.  
 / APPLICANT: Rank, David R.  
 / APPLICANT: Hanzel, David K.  
 / APPLICANT: Chen, Wenheng  
 / TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR MICROARRAY  
 / TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 / FILE REFERENCE: Aemica-X-1  
 / CURRENT APPLICATION NUMBER: US/09/864,761  
 / CURRENT FILING DATE: 2001-05-23  
 / PRIOR APPLICATION NUMBER: US 60/180,312  
 / PRIOR FILING DATE: 2000-02-04  
 / PRIOR APPLICATION NUMBER: US 60/207,456  
 / PRIOR FILING DATE: 2000-05-16  
 / PRIOR APPLICATION NUMBER: US 09/632,366  
 / PRIOR FILING DATE: 2000-08-03  
 / PRIOR APPLICATION NUMBER: GB 24263.6  
 / PRIOR FILING DATE: 2000-10-04  
 / PRIOR APPLICATION NUMBER: US 60/236,359  
 / PRIOR FILING DATE: 2000-09-27  
 / PRIOR APPLICATION NUMBER: PCT/US01/00666  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00667  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00664  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00669  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00665  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00668  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00663  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00662  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00661  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR APPLICATION NUMBER: PCT/US01/00670  
 / PRIOR FILING DATE: 2001-01-30  
 / PRIOR FILING DATE: 2000-09-21  
 / PRIOR FILING DATE: 2000-06-30  
 / PRIOR FILING DATE: 2001-01-29  
 / NUMBER OF SEQ ID NOS: 49117  
 / SOFTWARE: Aminoxam Sequence Listing Engine vers. 1.1  
 / SEQ ID NO: 15513  
 / LENGTH: 448  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8  
 / OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1  
 US-09-864-761-15513

Query Match 12.0%; Score 141.8; DB 9; Length 448;  
 Best Local Similarity 95.4%; Pred. No. 4.7e-33;  
 Matches 146; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 912 TGGTGGACATTACTCTGAGCTGGCGATGACATCTGCTTACTAAGAGCCGTG 971  
 QY 972 TCCTGAGAGGGCTGGCGCTTCTGCGAGATAACCCCTACCTGTGACTGTGCA 1031  
 Db 313 TCCTGAGAGGGCTGGCGCTTCTGCGAGATAACCCCTACCTGTGACTGTGCA 372

QY 912 TGGTGGACATTACTCTGAGCTGGCGATGACATCTGCTTACTAAGAGCCGTG 971  
 Db 269 TGGAAGTTCTCTTCTCAAGCTGGCGATGACATCTGCTTACTAAGAGCCGTG 328

QY 972 TCCTGAGAGGGCTGGCGCTTCTGCGAGATAACCCCTACCTGTGACTGTGCA 1031

Db 329 TCTGCAAGGGCTGGCGTCCCTGCAAGGATACCCCTACCTGTGACTGTGCAGA 388  
 Qy 1032 GGACCACTAACTGGAAGAGCTGAGACT 1064  
 Db 389 GGACCACTAACTGAAAGAGCTGAGACT 421

RESULT 11  
 US-09-814-353-4631  
 / Sequence 4631, Application US/0914353  
 / Publication No. US20030165831A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Lee, John  
 / APPLICANT: Thompson, Pamela  
 / APPLICANT: Lillie, James  
 / TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 / IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 / THERAPY OF OVARIAN CANCER  
 / FILE REFERENCE: MRI-06GB  
 / CURRENT APPLICATION NUMBER: US/09/814,353  
 / CURRENT FILING DATE: 2001-03-21  
 / PRIOR APPLICATION NUMBER: US 60/191,031  
 / PRIOR FILING DATE: 2000-03-21  
 / PRIOR APPLICATION NUMBER: US 60/207,124  
 / PRIOR FILING DATE: 2000-05-25  
 / PRIOR APPLICATION NUMBER: US 60/211,940  
 / PRIOR FILING DATE: 2000-06-15  
 / PRIOR APPLICATION NUMBER: US 60/216,820  
 / PRIOR FILING DATE: 2000-07-07  
 / PRIOR APPLICATION NUMBER: US 60/220,661  
 / PRIOR FILING DATE: 2000-07-25  
 / PRIOR APPLICATION NUMBER: US 60/257,672  
 / PRIOR FILING DATE: 2000-12-21  
 / NUMBER OF SEQ ID NOS: 22037  
 / SOFTWARE: FastSEQ for Windows Version 4.0  
 / SEQ ID NO: 4631  
 / LENGTH: 152  
 / OTHER INFORMATION: n = A,T,C or G  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE: misc\_feature  
 / NAME/KEY: misc\_feature  
 / LOCATION: 17, 102, 112  
 / OTHER INFORMATION: US-09-814-353-10930

Query Match Score 11.1%; DB 10; Length 152;  
 Best Local Similarity 95.0%; Pred. No. 4.8e-30;  
 Matches 134; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 536 GGCCTGGCGGAGCTGCTGAGACTGGGAGCATGACCATGGCTCTGAGATGGA 595  
 Db 12 GTGCGNGCCGGGGTCTGAGCTGGGAGCTGGGAGCATGGCTCTGAGATGGA 71  
 Qy 596 GACTGTGAGGGTGTGTTGAGCTGAGCTGGGAGGATAACATCCCAGCTCCAC 655  
 Db 72 GACTGTGAGGGTGTGTTGAGCTGAGCTGGGAGGATAACATCCCAGCTCCAC 131  
 Qy 656 GTGGGAAGAGCTCCATGG 676  
 Db 132 GTGCCCAAGTCCTCCATGG 152

RESULT 12  
 US-10-062-674-2038  
 / Sequence 2038, Application US/10062674  
 / Publication No. US20040005559A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.  
 / TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
 / FILE REFERENCE: PA-00026-1.CIP  
 / CURRENT APPLICATION NUMBER: US/10/062,674  
 / CURRENT FILING DATE: 2002-01-30  
 / PRIOR APPLICATION NUMBER: US 09/625,102  
 / PRIOR FILING DATE: 2000-07-24  
 / NUMBER OF SEQ ID NOS: 2217  
 / SOFTWARE: Perl Program  
 / SEQ ID NO: 2038  
 / LENGTH: 2343  
 / OTHER INFORMATION: Incyte ID No. US20040005559A1 411296.5  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE: misc\_feature  
 / NAME/KEY: misc\_feature  
 / LOCATION: 1-10-062-674-2038

Query Match Score 10.5%; DB 15; Length 2343;  
 Best Local Similarity 54.1%; Pred. No. 1.3e-27;  
 Matches 284; Conservative 0; Mismatches 239; Indels 12; Gaps 1;

Db 457 CCAAGGCCGGAGCTGCTGAGCTGGGAGCATGGCAAGCCAGAGGAGAAG 516

RESULT 13  
 US-10-062-674-2038  
 / Sequence 2038, Application US/10062674  
 / Publication No. US20040005559A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Lee, John  
 / APPLICANT: Thompson, Pamela  
 / APPLICANT: Lillie, James  
 / TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 / IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 / THERAPY OF OVARIAN CANCER  
 / TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 / IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 / THERAPY OF OVARIAN CANCER

Db	609	CAACGCAAACACCGGAAATCAGGGAGGGCTCTGAGCACATCATCGTGTGCGCT	6.8
Qy	517	GGGCAGTTCCGGGAGGTGCCGCCGAGCTGTGCTAGACTGGGAGGCAATTGAC	576
Db	669	GTATGATTACAGGGCATTACCCAGAAGACCTCGCTTCAAGGGGACCAGATGT	728
Qy	577	CATCTCTCTGAGGTGGAACTGTGGACGGGGTGTCGAAGTCTAGGGAGAGTA	636
Db	729	GGTCTTAGGAAATCGGGAGTGTGGAGGGCTGATCCCTGCAACCGGAGGG	788
Qy	637	TAACPTCCCAGCTCACTGGAAAATGTCCTCCAT-----GGTCTGCTGTA	684
Db	789	CTATCTCCAGAACATACTGTCGCCGGTGTGATTCTGGAGAACAGGGAGGT	848
Qy	685	TGAGGCGCTGGAGGAGGAGGAAAGCGAGGAAACTCTGGAAACCTGGAGG	744
Db	849	CAAGGGCATCGCCGAAGGACGGRAGGGCAACTGCTGGCTCCGGAAACATGCTGG	908
Qy	745	GCGCCCTCCATCGGGAGAACCTGTCGCCGGTCTACTCTCTGTCAGTCGGCT	804
Db	909	CTCCCTCATGATCGGGATAGCAGAACACTAAGGAACTACTCTGCTGGAGA	968
Qy	805	CAGCGCCCTGCACTCTGGACCTGAGACAGGATCACTGCACTGGATCACTGCCTGACAATGG	864
Db	969	CTACACCTCTGGAGGAGATACTGGAAACATTACAGATGGACCTGGACAAACGG	1028
Qy	865	CTGGCTGTACATCTACCCGGCTACCTCTCCCTACCTCAGGCCCTGGACACATTAA	924
Db	1029	GGGCTTCTACATATCCCCGAAAGCACCTTACGACTCTGAGGCTGGACACACTA	1088
Qy	925	CTCTGAGTGGGGATGAACTCTGCTGCTACTCAAGAGGCTCTGGACACCTG	969
Db	1089	CAAGAAGGGAAACGAGGGCTCTGCAGAAACTGTGCTGCTGCTG	1133
RESULT 14			
US-10-193-720-1			
; Sequence 1, Application US/10193720			
; Publication No. US20040009173A1			
; GENERAL INFORMATION:			
; APPLICANT: Friera, Annabelle M.			
; APPLICANT: Wong, Brian R.			
; APPLICANT: Masuda, Esteban			
; ATTORNEY: Powell, Mark			
; TITLE OF INVENTION: Modulators of Leukocyte Activation, HCK Compositions and			
; FILE REFERENCE: A-711313/RMS/DHR			
; CURRENT APPLICATION NUMBER: US/10/193,720			
; CURRENT FILING DATE: 2002-11-27			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: PatentIn version 3.1			
; SEQ_ID NO 1			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (78) ..(1595)			
; OTHER INFORMATION:			
US-10-193-720-1			
Query Match 10.4% Score 123; DB 15; Length 1924;			
Best Local Similarity 53.9%; Pred. No. 3.8e-27; Matches 283; Conservative 0; Mismatches 230; Indels 12; Gaps 1;			
Qy	457	CCAAGGCCAGGGACCTGTGACATGGGAGGCAAGGAGGCTGCGCTGGCCCT	516
Db	212	CAACAGAAACACAGGAAATGGGGGGGGCTTGAGGACATCTGGTGGCTGCC	271
Qy	517	GGGCAGTTCCGGGAGGTGCCGCCGAGCTGTGCTAGACTGGGAGGCTGAC	576
Db	272	GTATGATTACAGGGCATTACCAACAGGACCTCAGCTCCAGAAGGGGACCA	331
Qy	577	CATCGTCTCTAGGATGGAGACTGGCTGCAAGTCTGAGGAGAGCTGAGA	636

```

Db 332 GGTCTTAGGGATCCGGAGGGTGGAAAGGCTGACCCCTGCCAACGGAGGGG 3 91
Qy 637 TAACATCCCCAGGTCACGTGGCAAGTCTCCAT-----GGTGCGCTGTA 6 84
Db 392 CTACATCCAAGGAACATAGTCGCCCGGTTGACTCTGGAGAACAGGGAGCTT 4 51
Qy 685 TAGGGCTGAGGGAGAAGCAGGAACTGTGTGTTACCTGGGAAACCTGGAGG 7 44
Db 452 CAAGGGATCACCG3AGGGCAGGGCAACTGTGGCTCCGCCAACATGGGG 5 11
Qy 745 GGCCTTCCTCATCGGAGACCGAGGGCTTACTCTCTGAGCCGCTCT 8 04
Db 512 CTCCCTCATGATCCGGGATAGGAGAACCTAAGGAAGCTATCTGTGGAGAA 5 71
Qy 805 CAGCCGCCCTGCACTACAGGATCCAGACACTACAGGATCCACTGCCTTGACATGG 8 64
Db 572 CTAGGACCTCTGGACGGATCAAGCTACAGGATCCAGACACTACAGGATCC 8 64
Qy 865 CTGGCTCATCTCACCGGCTCACCTCCCTCTACTCCGGCCCTGGTGGACCTTA 9 24
Db 632 GGCTCTACATATCCCGAGAACCTCTGGAGGGCTGGAGGACCACTGGACCACTA 6 91
Qy 925 CTCTGAGCTGCCGATGACAATGCTGCTCTACTCAAGAGCCCTG 9 69
Db 692 CAAGAGGGAAACGACGGCTCTGCCORGAACACTGTCCGTGCCCTG 7 36

RESULT 15
US-09-954-456-1983
Sequence No. 1983, Application US/0954456
Patient No. US200211507A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using
TITLE OF INVENTION: Sets
FILE REFERENCE: 689220-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: Patentin version 3.0
SEQ ID NO: 1983
LENGTH: 2015
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1983

Query Match Score 10.4%; DB 9; Length 2015;
Best Local Similarity 53.9%; Pred. No. 3 9e-27;
Matches 283; Conservative 0; Mismatches 230; Indels 12; Gaps 1;
Qy 457 CCAGGCCAGGGACCTGACCATGGAAAGCAGAACGAAAGCCACAGCCGTTGGCCCT 5 16

```

Db	303	CAACAGCAACACCAAGGAAATAGGGAGGGAGGCTGTGAGGACATCATCGTGTGCCCT	362
Qy	517	GCGCAGTTCCGGCAGGTGGCCGGCGCGAAGCTGTGCTGAGACTGGGAGCCATTGAC	576
Db	363	GATGATAACGAGCCATTCAACCAAGAACCTCAACCTTCAGAGGGACCAAGATGCT	422
Qy	577	CATCGTCTCTGAGGATGGAGACTGGAGCTGGACGGTGTCTGAAGCTCTCGAGAGATA	636
Db	423	GTCCTAGAGGAAATCCGG3AGTGTTGAAAGCTGATCCCTGGCACCGGAGGAGG	482
Qy	637	TAACATCCCCAACCGTCCAGTGGCCAAGTTCCTA-----GGTGGCTGTA	684
Db	483	CTACATCCCCAACCAACTATGTGCCGGCGTGTACATCTCGAGACAGAGGAGTGT	542
Qy	685	TGAGGGCTGTGAGGGAGANAGAGAGGAACTGCTGTCTACCTGGAAACCTGGAGC	744
Db	543	CAAGGGCATCAGCGGAAGGICGAGAGCGCAACTGTGGCTCCGGAAATGCTGG	602
Qy	745	GCCCTCTCTATCGGGAGAACGCCAGACCAGAGGGCTTAACTCTGTGTCAGTCCGCCT	804
Db	603	CTCCCTCATGATCCGGATAACGGAGGAACTAAAGGAAGTACTTTTCGTCGAGA	662
Qy	805	CAGCCGCCTCATCTGGGCGGATCACTACAGGATCCACTGCTGCTGACAATGG	864
Db	663	CTAGGACCTCTGGAGGAGTACCGTGAACTAACAGATGACCTGGACCAACCG	722
Qy	865	CTGGCTGTACATCTACCGCGCCTCACCTTCCCCTCACTCAAGCCCTGGTGGACCATTA	924
Db	723	GGGTTCTACTATCCCCAAGGACCTTAGAACCTCTCGAGAGGCTGTGACCACTA	782
Qy	925	CTCTGAGCTGGGATGACATCTGTGCTCTACTAGGGACCCCTG	969
Db	783	CAAGAGGGAAACGAGCGGGCTCTGCAGAAAATGTCGTCGCTG	827

Search completed: February 20, 2004, 05:59:08  
 Job time : 2861.55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Run on:	February 20, 2004, 01:45:20 (without alignments)	Search time:	4572 Seconds
OM nucleic - nucleic search, using sw model			
Title:	US-03-939-853A-74	Minimum DB seq length:	0
Perfect score:	1183	Maximum DB seq length:	2000000000
Sequence:	1 agcttagactccaaaggaccc.....totctttggatgtgcctag 1183	Post-processing:	Listing first 45 summaries
Scoring table:	OLIGO NUC	Database :	GenEmbl*
Searched:	Gapop 60.0 , Gapext 60.0		1: gb_bt;*
Word size :	3470272 seqs, 21671516995 residues		2: gb_ng;*
	0		3: gb_ir;*
	Total number of hits satisfying chosen parameters: 6940544		4: gb_om;*
			5: gb_ov;*
			6: gb_bt;*
			7: gb_ph;*
			8: gb_dp;*
			9: gb_dr;*
			10: gb_ro;*
			11: gb_sts;*
			12: gb_sy;*
			13: gb_un;*
			14: gb_vl;*
			15: em_da;*
			16: em_fun;*
			17: em_hum;*
			18: em_in;*
			19: em_mu;*
			20: em_om;*
			21: em_or;*
			22: em_ov;*
			23: em_pat;*
			24: em_ph;*
			25: em_pl;*
			26: em_ro;*
			27: em_sts;*
			28: em_vl;*
			29: em_vn;*
			30: em_htg_hum;*
			31: em_htg_inv;*
			32: em_htg_other;*
			33: em_htg_mus;*
			34: em_htg_pln;*
			35: em_htg_rod;*
			36: em_htg_mam;*
			37: em_htg_vrt;*
			38: em_sy;*
			39: em_htgo_hum;*
			40: em_htgo_mus;*
			41: em_htgo_other;*

Pred. No. is the number of results predicted by chance to have a

RESULT	AX443133	LOCUS	AX443133	DEFINITION	Sequence 74 from Patent WO21699.
		ACCESSION	AX443133	VERSION	GI:21690555
		KEYWORDS	Homo sapiens (human)	SOURCE	Homo sapiens
		ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.	REFERENCE	Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R., Shimkets, R.A., Spytek, K.A., Szekeles, E.S., Tomlinson, J.B., Topper, J.N. and Yang, R.B.
		AUTHORS			

TITLE	Proteins and nucleic acids encoding same											
JOURNAL	Patent: WO 0216599-A 74 28-FEB-2002;											
FEATURES	Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)											
source	1..1183 /organism="Homo sapiens" /mol type="unassigned DNA" /db_xref="taxon:9606"											
ORIGIN												
Query Match	Score 1183 ; DB 6 ; Length 1183 ;											
Best Local Similarity	100.0% ; Pred. No. 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;											
Matches 1183 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;	Qy	1	AGCTAGAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Db	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Qy	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Db	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60
Qy	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Db	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Qy	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Db	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	
Qy	61	CCTTCCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 60	Db	61	CCTTCCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 60	Qy	61	CCTTCCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 60	Db	61	CCTTCCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 60	
Qy	121	CATGGGCCAGTGTCATCCCTGGTGAAACTCTGACTGCAGACATGGCTG 180	Db	121	CATGGGCCAGTGTCATCCCTGGTGAAACTCTGACTGCAGACATGGCTG 180	Qy	121	CATGGGCCAGTGTCATCCCTGGTGAAACTCTGACTGCAGACATGGCTG 180	Db	121	CATGGGCCAGTGTCATCCCTGGTGAAACTCTGACTGCAGACATGGCTG 180	
Qy	181	ACCCAAACCAACACCTAGCTCCTCCCTGAAGATCTCCCTGGTGAAACTCTG 180	Db	181	ACCCAAACCAACACCTAGCTCCTCCCTGAAGATCTCCCTGGTGAAACTCTG 180	Qy	181	ACCCAAACCAACACCTAGCTCCTCCCTGAAGATCTCCCTGGTGAAACTCTG 180	Db	181	ACCCAAACCAACACCTAGCTCCTCCCTGAAGATCTCCCTGGTGAAACTCTG 180	
Qy	241	TCTCTAGACCAAGGACACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 240	Db	241	TCTCTAGACCAAGGACACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 240	Qy	241	TCTCTAGACCAAGGACACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 240	Db	241	TCTCTAGACCAAGGACACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 240	
Qy	301	GCCAGAGCATCGCTCCAGAGAGCTCCAGACTCCAGCTCCAGCTCCAGCT 360	Db	301	GCCAGAGCATCGCTCCAGAGAGCTCCAGACTCCAGCTCCAGCTCCAGCT 360	Qy	301	GCCAGAGCATCGCTCCAGAGAGCTCCAGACTCCAGCTCCAGCTCCAGCT 360	Db	301	GCCAGAGCATCGCTCCAGAGAGCTCCAGACTCCAGCTCCAGCTCCAGCT 360	
Qy	361	CTCGATGATGCTGCTCTGAGTCGCTGCTGAGTCGAGAATGGAACTGCTG 420	Db	361	CTCGATGATGCTGCTCTGAGTCGCTGCTGAGTCGAGAATGGAACTGCTG 420	Qy	361	CTCGATGATGCTGCTCTGAGTCGCTGCTGAGTCGAGAATGGAACTGCTG 420	Db	361	CTCGATGATGCTGCTCTGAGTCGCTGCTGAGTCGAGAATGGAACTGCTG 420	
Qy	421	AAAATCTCTGCGCAAGCCAGCTGAGTCGCTGAGTCGCTGAGTCGCTG 480	Db	421	AAAATCTCTGCGCAAGCCAGCTGAGTCGCTGAGTCGCTGAGTCGCTG 480	Qy	421	AAAATCTCTGCGCAAGCCAGCTGAGTCGCTGAGTCGCTGAGTCGCTG 480	Db	421	AAAATCTCTGCGCAAGCCAGCTGAGTCGCTGAGTCGCTGAGTCGCTG 480	
Qy	481	GGAGAGAGAGAAAGCAAGGCCCAAGGCCAGGGCTGGCCAGTTCGGCC 540	Db	481	GGAGAGAGAGAAAGCAAGGCCCAAGGCCAGGGCTGGCCAGTTCGGCC 540	Qy	481	GGAGAGAGAGAAAGCAAGGCCCAAGGCCAGGGCTGGCCAGTTCGGCC 540	Db	481	GGAGAGAGAGAAAGCAAGGCCCAAGGCCAGGGCTGGCCAGTTCGGCC 540	
Qy	541	GGCCSAGAGCTGTCTGCTGAGACTCGGGAGAGCATTGCACTGCTGGAGACTG 600	Db	541	GGCCSAGAGCTGTCTGCTGAGACTCGGGAGAGCATTGCACTGCTGGAGACTG 600	Qy	541	GGCCSAGAGCTGTCTGCTGAGACTCGGGAGAGCATTGCACTGCTGGAGACTG 600	Db	541	GGCCSAGAGCTGTCTGCTGAGACTCGGGAGAGCATTGCACTGCTGGAGACTG 600	
Qy	601	GTGGAGCTCCATGGGTCTGTGAGAGTATAACATCCCAGGCTCACGTGG 660	Db	601	GTGGAGCTCCATGGGTCTGTGAGAGTATAACATCCCAGGCTCACGTGG 660	Qy	601	GTGGAGCTCCATGGGTCTGTGAGAGTATAACATCCCAGGCTCACGTGG 660	Db	601	GTGGAGCTCCATGGGTCTGTGAGAGTATAACATCCCAGGCTCACGTGG 660	
Qy	661	CAAAGCTCCATGGGTCTGTGAGAGTATAACATCCCAGGCTCACGTGG 720	Db	661	CAAAGCTCCATGGGTCTGTGAGAGTATAACATCCCAGGCTCACGTGG 720	Qy	661	CAAAGCTCCATGGGTCTGTGAGAGTATAACATCCCAGGCTCACGTGG 720	Db	661	CAAAGCTCCATGGGTCTGTGAGAGTATAACATCCCAGGCTCACGTGG 720	
Qy	721	GTGGTTACCTGGAAACCCCTGGAGGGCTTCTCTGAGTCGAGTCGAGTCG 780	Db	721	GTGGTTACCTGGAAACCCCTGGAGGGCTTCTCTGAGTCGAGTCGAGTCG 780	Qy	721	GTGGTTACCTGGAAACCCCTGGAGGGCTTCTCTGAGTCGAGTCGAGTCG 780	Db	721	GTGGTTACCTGGAAACCCCTGGAGGGCTTCTCTGAGTCGAGTCGAGTCG 780	
Qy	781	CTCTTACTCTGTGCTGAGTCGCTGCTGAGTCGAGTCGAGTCGAGTCG 840	Db	781	CTCTTACTCTGTGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCGAGTCG 840	Qy	781	CTCTTACTCTGTGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCGAGTCG 840	Db	781	CTCTTACTCTGTGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCGAGTCG 840	
Qy	841	CAGATCCACTGTGCTGAGTCGCTGAGTCGCTGAGTCGAGTCGAGTCG 900	Db	841	CAGATCCACTGTGCTGAGTCGCTGAGTCGCTGAGTCGAGTCGAGTCG 900	Qy	841	CAGATCCACTGTGCTGAGTCGCTGAGTCGCTGAGTCGAGTCGAGTCG 900	Db	841	CAGATCCACTGTGCTGAGTCGCTGAGTCGCTGAGTCGAGTCGAGTCG 900	
Qy	901	ACTCCAGGCCCTGGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCG 960	Db	901	ACTCCAGGCCCTGGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCG 960	Qy	901	ACTCCAGGCCCTGGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCG 960	Db	901	ACTCCAGGCCCTGGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCG 960	
Qy	961	GGAGCCCTGGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCGAGTCG 1020	Db	961	GGAGCCCTGGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCGAGTCG 1020	Qy	961	GGAGCCCTGGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCGAGTCG 1020	Db	961	GGAGCCCTGGCTGAGTCGCTGAGTCGAGTCGAGTCGAGTCGAGTCG 1020	
Qy	1021	GACTGTGAGAGCAACACTGAGTCGAGTCGAGTCGAGTCGAGTCG 1080	Db	1021	GACTGTGAGAGCAACACTGAGTCGAGTCGAGTCGAGTCGAGTCG 1080	Qy	1021	GACTGTGAGAGCAACACTGAGTCGAGTCGAGTCGAGTCGAGTCG 1080	Db	1021	GACTGTGAGAGCAACACTGAGTCGAGTCGAGTCGAGTCGAGTCG 1080	
Qy	1081	AGCTGCACAGGGAGAGCTCTCTGAGTCGAGTCGAGTCGAGTCG 1140	Db	1081	AGCTGCACAGGGAGAGCTCTCTGAGTCGAGTCGAGTCGAGTCG 1140	Qy	1081	AGCTGCACAGGGAGAGCTCTCTGAGTCGAGTCGAGTCGAGTCG 1140	Db	1081	AGCTGCACAGGGAGAGCTCTCTGAGTCGAGTCGAGTCGAGTCG 1140	
Qy	1141	CATCAGGCTGAATGAGTCGAGTCGAGTCGAGTCGAGTCG 1183	Db	1141	CATCAGGCTGAATGAGTCGAGTCGAGTCGAGTCGAGTCG 1183	Qy	1141	CATCAGGCTGAATGAGTCGAGTCGAGTCGAGTCGAGTCG 1183	Db	1141	CATCAGGCTGAATGAGTCGAGTCGAGTCGAGTCGAGTCG 1183	
RESULT 2												
LOCUS	AX443135 /c											
DEFINITION	Sequence 76 from Patent WO0216599.											
VERSION	AX443135											
KEYWORDS	Homo sapiens (human)											
ORGANISM	Homo sapiens											
MATERIAL	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Primates; Catarrhini; Hominidae; Homo.											
REFERENCE	1 Burgess,C.E., Conley,P.B., Grossie,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spyrek,K.A., Szelekes,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.											
AUTHORS	Burgess,C.E., Conley,P.B., Grossie,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spyrek,K.A., Szelekes,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.											
TITLE	Proteins and nucleic acids encoding same											
JOURNAL	Patent: WO 0216599-A 76 28-FEB-2002;											
FEATURES	Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)											
source	1..1183 /organism="Homo sapiens" /mol type="unassigned DNA" /db_xref="taxon:9606"											
ORIGIN												
Query Match	Score 1183 ; DB 6 ; Length 1183 ;											
Best Local Similarity	100.0% ; Pred. No. 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;											
Matches 1183 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;	Qy	1	AGCTAGAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Db	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Qy	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60	Db	1	AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 60
Qy	61	CCTTCCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 60	Db	61	CCTTCCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 60	Qy	61	CCTTCCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 60	Db	61	CCTTCCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 60	
Qy	1183	1..1183	1..1183	1..1183	1..1183	1..1183	1..1183	1..1183	1..1183	1..1183		
Qy	1183	/organism="Homo sapiens"	0	0	0	0	0	0	0	0		
Db	1183	/mol type="unassigned DNA"	0	0	0	0	0	0	0	0		
Qy	1183	/db_xref="taxon:9606"	0	0	0	0	0	0	0	0		
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											
Query	1 AGCTAGCTCCAAGGACCCAGCCTGCTGTGAAGAGTCAAAGGCCCTGCG 960											



Db	858	CAGATCCACTGCTTACAATGGTGGTACATCAGCGCTCCCTCCCC 917	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www-sigc.stanford.edu">http://www-sigc.stanford.edu</a>
Qy	901	ACTCCAGCCCTGGACATTAATCTGAGTTGGGATGACATCTGGCCCTAA 960	Contact: (Dickson, Mark) <a href="mailto:mcod@paxil.stanford.edu">mcod@paxil.stanford.edu</a>
Db	918	ACTCCAGCCCTGGTGTGGACATTAATCTGAGTTGGGATGACATCTGGCCCTAA 977	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Qy	961	GGAGCCCTGTGTGCTTGAGAGGCTGCCCTCCCTGGCAAGGTTACCCCTAACCTGT 1020	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium (LNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>
Db	978	GGAGCCCTGTGTGCTTGAGAGGCTGCCCTCCCTGGCAAGGTTACCCCTAACCTGT 1037	Series: IFAK plate: 88 Row: a Column: 20 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28416422.
Qy	1021	GACTGTCAGAGGACACCACTCAACTGGAAAGAGCTGGAAGAGCTGCCCTGGCAAGGTTACCCCTAACCTGT 1080	Location/Qualifiers
Db	1038	GACTGTCAGAGGACACCACTCAACTGGAAAGAGCTGGAAGAGCTGCCCTGGCAAGGTTACCCCTAACCTGT 1097	1. 2538 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:43845 IMAGE:4429896" /tissue_type="Prostate, adenocarcinoma." clone_id="NIH MGC_91" /lab_host="DH10B" /note="Vector: PCMV-SPORT6"
Qy	1081	AGCTGCCACAGGGAGAGGCTCTCTCATGGAGGCTCTGGGAGTCAGCTCTA 1140	1. 2538 /gene="SLC22" /note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845, SLAP2" /db_xref="LocusID:94174" /db_xref="MMI:606577" /db_xref="LocusID:84174" /product_start=1 /product_id="PAAH42041_1" /protein_id="PAAH42041_1" /db_xref="GI:127468943" /db_xref="LocusID:84174" /translation=MGSIPSPRSLSPLPSLSSVQOGPVTMEAEERKATAVALGSFP AGGPAAEISURGEPLTIVSEEDDMWITVSESGRENNPISWDEIRYRTHCLDNGWLX KAEEFLIIPNPGGAFLRESQRGRGTSLSVRSLRSPWDLRHYRTHCLDNGWLX SPRTPFSPLQVQASPLPQDIPLPVTVORTPLNK ELDSLLSSEAATGEESLSEIRESLSFYISUNDEAVSLDDA"
Db	1098	AGCTGCCACAGGGAGAGGCTCTCTCATGGAGGCTCTGGGAGTCAGCTCTA 1157	1. 2538 /note="SH3: Region: SH3 domain. SH3 (Src homology 3 domains are often indicative of a protein involved in signal transduction related to cytoskeletal organization. First described in the Src cytoplasmic tyrosine kinase. The structure is a partly opened beta barrel" /db_xref="CD:pfam00018" /note="SH2: Region: SH2 domain" /db_xref="CD:pfam00017"
Qy	1141	CATCAGCTGTATGATGAGGCTGCTCTGGATGATCCCTAG 1183	ORIGIN
Db	1158	CATCAGCTGTATGATGAGGCTGCTCTGGATGATCCCTAG 1200	Query Match 88.4%; Score 1046; DB 9; Length 2538; Best Local Similarity 99.8%; Pred. No: 0; Gaps 0; Matches 1146; Conservative 0; MisMatches 2; Indels 0;
RESULT 4			Qy 36 GTGACAGAGCTCAAAGGCCCTGGCTGCTGGCTGGAGGGT 95 Db 1 GTGACAGAGCTCAAAGGCCCTGGCTGCTGGAGGGT 60
LOCUS	BC042041	2538 bp mRNA linear PRI 07-OCT-2003	Qy 96 TCCCCAGTCAGATCCCTAAGGAGCTGGGAGGCTCATCCCTGGTACAACT 155 Db 61 TCCCCAGTCAGATCCCTAAGGAGCTGGGAGGCTCATCCCTGGTACAACT 120
DEFINITION		Clone MGC:43845 IMAGE:4429896, complete cds.	Qy 156 GCTGACTGGAGACAGATGCTGAGCTTACCCAACCAACCTGGCTGGCTGGAGATCC 215 Db 121 GCTGACTGGAGACAGATGCTGAGCTTACCCAACCAACCTGGCTGGAGATCC 180
ACCESSION	BC042041		Qy 216 TCCCAAGGCTGAGGAGCTCTGGCTGCTTAGGCCAGACCTGGAGACTTCAGAG 275 Db 181 TCCCAAGGCTGAGGAGCTCTGGCTGCTTAGGCCAGACCTGGAGACTTCAGAG 240
VERSION	BC042041.1		Qy 276 GGCCOCCTAAAGCCCTAACCTGTCAGCCAGAGATGCCTAGCACAGCTTCAGCA 335
KEYWORDS	MGC,		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.		
REFERENCE	1 (bases 1 to 2538)		
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Sheinin, C.M., Schuler, G.D., Altenschul, S.F., Zeeberg, B., Buetow, K.H., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Dapchenko, I., Maruska, K., Farmer, A.A., Rubin, G.M., Hong, L., Sharpston, M., Soares, M.B., Brownstein, M.J., Casavant, T.L., Scheetz, T.E., Prange, C., Raha, S., S. Loquellano, N.A., Peters, G.J., Carninci, P., Mullany, S. J., Bosak, S.A., McEvian, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalon, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, R. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., M., Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smailis, D. B., Schnerch, A., Schein, J. E., Jones, S. J., and Marrs, M. A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) PubMed 12477287		
REFERENCE	2 (bases 1 to 2538)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-DIC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
COMMENT	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
REMARK	Contact: mgc-help@mail.nih.gov		
COMMENT	Tissue Procurement: DCID-DTP		
CDNA Library Preparation: Life Technologies, Inc.			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			











ACCESSION	AF290986	Qy	878	TCACCGCCCTACCTTCCCTCACTGGAAAGCTGTTGACCTTACTGTGAG
VERSION	GI:17351922	Db	481	TCACCGCCCTACCTTACTGGAAAGCTGTTGACCTTACTGTGAG
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	Loreto, M.P. and McGlade, C.J.	RESULT	13	
AUTHORS		HS460J8/C		
TITLE	Cloning and characterization of human Src-like adaptor protein 2 and a novel splice isoform, SRAp-2-v	LOCUS	HS460J8	66741 bp DNA linear
JOURNAL	Oncogene 22 (2), 266-273 (2003)	DEFINITION		PRI 23-JUL-2001
MEDLINE				Human DNA sequence from clone RP3-460J8 on chromosome 20q11.21-11.23 Contains the 3' end of the gene for a novel protein similar to N-myc downstream regulated (NDRG1) the 5' end of a gene encoding a novel protein tyrosine kinase, ESTs, STSs and GSSs, complete sequence.
PMID	12557895	ACCESSION	AL031662	AL031662_26 GI:9716901 RTG; NDRG1; SH2 domain.
REVERSE	2 (bases 1 to 737)	VERSION	AL031662_26 GI:9716901	
AUTHORS	Loreto, M.P. and McGlade, C.J.	KEYWORDS		
TITLE	Direct Submission	SOURCE		
JOURNAL	Submitted (23-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada	ORGANISM		
FEATURES	Location/Qualifiers	REFERENCE		
source	1..737 /organism="Homo sapiens"	AUTHORS	Skuce,C.	
CDS	/mol_type="mRNA"	TITLE	Direct Submission	
	/db_xref="taxon:2606"	JOURNAL	(23-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk	
	/protein_id=AAI38198.1"	COMMENT	On Aug 7, 2000 this sequence version replaced gi:645549.	
	/db_xref="GI:117351923"		During sequence assembly data are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.	
	/translation="MGSPLSPRKLSPSLSSVQOGPVTEAERSKATAVALGSFP AGGAENISRLGPVITVEDDDNNNTVISESGRENNIPSYVHARVSHNLYLSRSRKAEPLNPGPAAFLRESQRRGKPLPSLSPASWIRHVRIRKCLDNLQWLYI SPRLTPSPSQALVTHYSEGWPAPWQGYTPTCDADTQLERAQOLPPF"		The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:	
ORIGIN	codon_start1		Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at	
	/product="Src-like adaptor protein-2 splice isoform"		http://www.sanger.ac.uk/projects/C-elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at	
	/protein_id=AAI38198.1"		http://www.sanger.ac.uk/HPG/Chr20	
	/db_xref="GI:117351923"		IMPORTANT: This sequence is not the entire insert of clone RP3-460J8 at 66741 in this sequence.	
	/transl_start1		It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.	
	/product="Src-like adaptor protein-2 splice isoform"		The true right end of clone RP3-460J8 is at 41767 in this sequence.	
	/protein_id=AAI38198.1"		The true left end of clone RP3-460J8 is at 41767 in this sequence.	
	/db_xref="GI:117351923"		This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., paired quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP3-460J8 is from the library RP01-3 constructed by the group of Pieter de Jong. For further details see	
			http://www.chori.org/bacpac/home.htm	
Query Match	40.8% Score 483; DB 9; Length 737;	FEATURES	VECTOR: pXPAC2	
Best Local Similarity	99.8%; Pred. No. 1..9..268;	source	Location/Qualifiers	
Matches	533; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		1..66741	
Qy	398 ATGGGAGTGTGCCAGCAAGAAATCTCTGCAGCTGAGTTCTCTGTC		/organism="Homo sapiens"	
Db	1 ATGGGAGTGTGCCAGCAAGAAATCTCTGCAGCTGAGTTCTCTGTC		/mol_type="genomic DNA"	
Qy	458 CAAGGCCAGGACACTGTGACATGGAAAGGAGAAAGGAGAAAGCCTG		/db_xref="Taxon:9606"	
Db	61 CAAGGCCAGGACACTGTGACATGGAAAGGAGAAAGGAGAAAGCCTG		/chromosome="20"	
Qy	518 GCGAGTTCCGGCAAGTGGCCGCACGCTGTCGAACTCGGGAGAAC		/map="q11.21-11.23"	
Db	518 GCGAGTTCCGGCAAGTGGCCGCACGCTGTCGAACTCGGGAGAAC		/clone="RP3-460J8"	
Qy	121 GCGAGTTCCGGCAAGTGGCCGCACGCTGTCGAACTCGGGAGAAC		/clone_1ib="RPCI-3"	
Db	121 GCGAGTTCCGGCAAGTGGCCGCACGCTGTCGAACTCGGGAGAAC		complement 50..544)	
Qy	578 ATGGTCCTGAGGATGGAGCTGAGTCTGAGCTGAGGAGAGAT		/note="match": GSS: B45150"	
Db	181 ATGGTCCTGAGGATGGAGCTGAGTCTGAGCTGAGGAGAGAT		complement 240..7128)	
Qy	638 AACATCCCACGTGTCACGTGGCAAGCTCCATGCTGTAGGGCCCTGAGC		/gene="ndu977B1.1"	
Db	241 AACATCCCACGTGTCACGTGGCAAGCTCCATGCTGTAGGGCCCTGAGC			
Qy	698 AGGGAGAAAGCAGAGGAACGCTGAGCTGAGCTGAGCTGAGCTGAGC			
Db	301 AGGGAGAAAGCAGAGGAACGCTGAGCTGAGCTGAGCTGAGCTGAGC			
Qy	758 CGGAGAGCGACGAGGAAGGGCTTACTCTGTCAGTCGGCTCCCTGCA			
Db	361 CGGAGAGCGACGAGGAAGGGCTTACTCTGTCAGTCGGCTCCCTGCA			
Qy	818 TCCTGGGACCGATCAGACTACAGGATCACCTGAGCATGCTGTGACATC			
Db	421 TCCTGGGACCGATCAGACTACAGGATCACCTGAGCATGCTGTGACATC			

mRNA complement[join(<240. .339.6995. .7128)]  
 /gene="dj9781.1"  
 /product="dj460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"  
 /note="match: cDNAs: Em:AK025645  
 match: ESTs: Em:BG178187"  
 /evidence="not experimental"  
 complement[join(<240. .339.6995. .7085)]  
 /gene="dj9781.1"  
 /note="continues in Em:AL050318 as dj977B1.2"  
 /codon\_start=1  
 /evidence="not experimental"  
 /product="dj460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"  
 /note="match: cDNAs: Em:AK025645  
 /protein\_id="CAC44645.1"  
 /db\_xref="GI:15020830"  
 /db\_xref="GOA:Q9H6Q3"  
 /db\_xref="SWISS-PROT:Q9H6Q3"  
 /translation="MGSILSPRSRKSLPSPLSSSVQQGPVTEAERSKATAVALGSFP  
 AGGAEULSLRGLPFLIVS"  
 536 . .565  
 /note="15 copies 2 mer ag 93% conserved"  
 2175 . .2402  
 /note="14 copies 2 mer ta 100% conserved"  
 2783 . .3155  
 /note="match: GSS: Em:AQ887191"  
 misc\_feature complement[6983 . .7482]  
 /note="match: GSS: Em:AQ556467"  
 misc\_feature complement[6985 . .7492]  
 /note="match: GSS: Em:7925  
 misc\_feature complement[6985 . .7492]  
 /note="match: STS: Em:HS427J1S"  
 repeat\_region 10615 . .10640  
 /note="13 copies 2 mer tt 92% conserved"  
 repeat\_region 13593 . .13661  
 /note="MER21B repeat: matches 422. .785 of consensus"  
 14248 . .14297  
 repeat\_region 14594 . .14614  
 /note="MER21B repeat: matches 374. .422 of consensus"  
 repeat\_region 14746 . .15078  
 /note="MER21B repeat: matches 355. .374 of consensus"  
 17519 . .17686  
 complement[17516]  
 /note="match: GSS: Em:AZ067993"  
 misc\_feature complement[17521]  
 polyA\_site complement[17521]  
 polyA\_site complement[17522]  
 gene complement[17523] . .54534]  
 /gene="dj469A13.3"  
 mRNA complement[join(<17523 . .19451,20551 . .22110, . .22161,  
 22247 . .22282,26087 . .26134,30796 . .30847,30986 . .31051,  
 32039 . .32142,37091 . .37147,46556 . .46552,48268 . .48328,  
 50163 . .50225,53242 . .53362,54429 . .>54534)]  
 /geno="dj469A13.3"  
 /product="dj460J8.1 (continued from dj469A13.3 in  
 Em:AL132768)"  
 /note="match: cDNAs: Em:AP03922 Em:U52073 Em:D87953  
 Em:AP159092 Em:AP147402 Em:M5814  
 matc: ESTs: Em:AA178226 Em:AU03952 Em:W89263  
 Em:AA325826 Em:AA120982 Em:AA162260 Em:AA445016  
 Em:AV002395 Em:T85147 Em:AI186615 Em:AI112437 Em:AI004026  
 Em:TP705 Em:AT786673 Em:AI181197 Em:AA764653 Em:AA316771  
 Em:AO035165 Em:AV002368"  
 /evidence="not experimental"  
 complement[17522 . .17547]  
 /geno="dj469A13.3"  
 17555 . .17811  
 misc\_feature /note="match: STS: Em:G19945"  
 join(1920 . .19451,2210 . .22282,26087 . .26127,30796 . .30847)  
 /note="match: STS: Em:c23762"  
 complement[join(19270 . .19451,20551 . .20599,22110 . .22161,  
 22247 . .22282,26087 . .26134,30796 . .30847,30986 . .31051,  
 32039 . .32142,37091 . .37147,46556 . .46552,48268 . .48328,  
 50163 . .50225,53242 . .53362,54429 . .>54534)]  
 /evidence="not experimental"  
 complement[17523 . .2541]  
 /geno="dj469A13.3"  
 17556 . .17811  
 misc\_feature /note="match: STS: Em:G19945"  
 join(1920 . .19451,2210 . .22282,26087 . .26127,30796 . .30847)  
 /note="match: STS: Em:c23762"  
 complement[join(19270 . .19451,20551 . .20599,22110 . .22161,  
 22247 . .22282,26087 . .26134,30796 . .30847,30986 . .31051,  
 32039 . .32142,37091 . .37147,46556 . .46552,48268 . .48328,  
 50163 . .50225,53242 . .53362,54429 . .>54534)]  
 /evidence="not experimental"

RESULT 14  
 AC026539

LOCUS 145833 bp DNA linear HTG 27-APR-2000  
 DEFINITION Homo sapiens chromosome 20 clone RP1-712N14 map 20, WORKING DRAFT  
 SEQUENCE, 37 unorderered pieces.

ACCESSION AC026539

VERSION 2 GI:7656813

KEYWORDS HTG; PHASEL; HTGS; DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE (bases 1 to 145833)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C., and Lander,E.  
 TITLE Homo sapiens chromosome 20, clone RP1-712N14  
 JOURNAL Unpublished  
 JOURNAL 2 (bases 1 to 145833)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barra,N., Bastien,V., Beda,F.,  
 Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,  
 Campopiano,A., Castle,A., Choepe,L.Y., Collangello,M., Collins,S.,  
 Collimore,A., Cooke,P., DeAngelis,K., Diaz,J.S.,  
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
 Gallegan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
 Grand-Pierre,N., Grant,G., Haggos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
 Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lebocky,J.,  
 Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
 Medlar,J., Meneus,J.L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
 Murphy,T., Nayef,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,  
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schaefer,S., Severy,P., Spencer,B.,  
 Strange-N., Stoyanovic,N., Subramanian,A., Talamas,J.,  
 Testray,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

JOURNAL Direct Submission  
 COMMENT Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 27, 2000 this sequence version replaced gi:7283243.  
 All repeats were identified using RepeatMasker:  
 Smit,A.F.A. & Green,P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

JOURNAL Genome Center  
 COMMENT Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Center Project name: L7115  
 Center Clone name: 712\_N14  
 Center Statistics

Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-Terminator Big Dye; 100% of reads  
 Assembly program: Phrap; Version 0.960731  
 Consensus quality: 1125577 bases at least Q40  
 Consensus quality: 1135793 bases at least Q30  
 Consensus quality: 1139593 bases at least Q20  
 Insert size: 1422233; agarose-fp  
 Quality coverage: .6 in Q20 bases; agarose-fp  
 Quality coverage: 3.3 in Q20 bases; sum-of-contigs  
 Quality coverage: 3.3 in Q20 bases; sum-of-contigs  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 37 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 1214: contig of 1214 bp in length  
 \* 1215 1314: gap of 100 bp  
 \* 1315 2673: contig of 1359 bp in length  
 \* 2674 2773: gap of 100 bp  
 \* 2774 4520: contig of 1747 bp in length  
 \* 4521 4620: gap of 100 bp  
 \* 4621 5961: contig of 1341 bp in length  
 \* 5962 6061: gap of 100 bp  
 \* 6062 7719: contig of 1658 bp in length  
 \* 7720 7819: gap of 100 bp  
 \* 7820 9799: contig of 1980 bp in length  
 \* 9800 9899: gap of 100 bp  
 \* 9900 11434: contig of 1535 bp in length  
 \* 11435 11534: gap of 100 bp  
 \* 11535 14382: contig of 2848 bp in length  
 \* 14383 14482: gap of 100 bp  
 \* 14483 17148: contig of 2666 bp in length  
 \* 17149 17248: gap of 100 bp  
 \* 17249 19239: contig of 1931 bp in length  
 \* 19240 19339: gap of 100 bp  
 \* 19340 21102: contig of 1763 bp in length  
 \* 21103 21204: gap of 100 bp  
 \* 21204 23371: contig of 2169 bp in length  
 \* 23372 23471: gap of 100 bp  
 \* 23472 25782: contig of 211 bp in length  
 \* 25783 25882: gap of 100 bp  
 \* 25883 28824: contig of 2942 bp in length  
 \* 28825 28924: gap of 100 bp  
 \* 28925 31611: contig of 2995 bp in length  
 \* 31620 31719: gap of 100 bp  
 \* 31720 34679: contig of 2960 bp in length  
 \* 34680 34779: gap of 100 bp  
 \* 34780 34791: contig of 2748 bp in length  
 \* 34792 37621: gap of 100 bp  
 \* 37628 40146: contig of 2519 bp in length  
 \* 40147 40244: gap of 100 bp  
 \* 40247 43743: contig of 3497 bp in length  
 \* 43743 43843: gap of 100 bp  
 \* 43844 47133: contig of 2920 bp in length  
 \* 47134 47233: gap of 100 bp  
 \* 47234 51023: contig of 3720 bp in length  
 \* 51024 51123: gap of 100 bp  
 \* 51124 54935: contig of 3812 bp in length  
 \* 54936 55035: gap of 100 bp  
 \* 55036 59532: contig of 4518 bp in length  
 \* 59532 59654: gap of 100 bp  
 \* 59654 64428: contig of 4775 bp in length  
 \* 64429 64528: gap of 100 bp  
 \* 64529 69211: contig of 4683 bp in length  
 \* 69212 69311: gap of 100 bp  
 \* 69312 72901: contig of 3590 bp in length  
 \* 72902 73002: gap of 100 bp  
 \* 73002 76724: contig of 3723 bp in length  
 \* 76725 76825: gap of 100 bp

FEATURES

source

misc\_feature

81179: contig of 4355 bp in length  
 81180: gap of 100 bp  
 81280: contig of 5730 bp in length  
 \* 87010: gap of 100 bp  
 \* 87110: contig of 3746 bp in length  
 \* 90855: gap of 100 bp  
 \* 90955: contig of 5565 bp in length  
 \* 96521: gap of 100 bp  
 \* 96621: contig of 5701 bp in length  
 \* 102321: gap of 100 bp  
 \* 102421: contig of 5872 bp in length  
 \* 108233: gap of 100 bp  
 \* 108294: contig of 896 bp in length  
 \* 116690: gap of 100 bp  
 \* 116739: contig of 8475 bp in length  
 \* 125255: gap of 100 bp  
 \* 125354: contig of 10990 bp in length  
 \* 136355: gap of 100 bp  
 \* 136455: contig of 9379 bp in length.

Location/Qualifiers

1..145833

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20"  
 /clone\_id="RP11-712N14"  
 /clone\_lib="RPCI-11 Human Male BAC"

misc\_feature

1..1214

/note="assembly\_fragment"  
 1315..2673

misc\_feature

2774..4520

misc\_feature

4622..5161

misc\_feature

6062..7719

misc\_feature

7820..9799

misc\_feature

9900..11434

misc\_feature

1155..14382

misc\_feature

14433..17148

misc\_feature

17249..19239

misc\_feature

19340..21102

misc\_feature

21203..23371

misc\_feature

23442..25782

misc\_feature

25833..28824

misc\_feature

28945..31619

misc\_feature

31700..34679

misc\_feature

34700..34757

misc\_feature

37658..40146

misc\_feature

40247..43743

misc\_feature

43844..47133

misc\_feature

47244..51023

misc\_feature

51144..54935

clone\_end:T7

vector\_side:right"  
 misc\_feature 55036..59553

Query Match 29.9% Score 354; DB 2; Length 145833;  
 Best Local Similarity 100.0%; Pred. No. 1.e-193;  
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCTAGAGTCAGGACCCAGGAACTTGTGTCAGACAGCTCAAAGGCCCTGGG 60  
 db 99745 AGCTAGAGTCAGGACCCAGGAACTTGTGTCAGACAGCTCAAAGGCCCTGGG 99804  
 Qy 61 CCTTCCCCTCCCTGCGCTTGAGGGTTCCGCTTGAGCTGAACTCTAAGAG 120  
 db 99805 CCTTCCCCTCCCTGCGCTTGAGGGTTCCGCTTGAGCTGAACTCTAAGAG 99864  
 Qy 121 CATGGGGCTCATCCATCCCTGGTTGAACACTGTGACTGCAGACAGATGGCTGAGCT 180  
 db 99865 CATGGGGCTGACTGATCCATCCCTGGTTGAACACTGTGACTGCAGACAGATGGCTGAGCT 99924  
 Qy 181 ACCCAAAACACACCTACCTGCTCCCTGAGATCCTCCAGGCTGAGAGTTCTGGTG 240  
 db 99925 ACCCAAAACACACCTACCTGCTCCCTGAGATCCTCCAGGCTGAGAGTTCTGGTG 99984  
 Qy 241 TCCPAGGACCAAGGACACTGGAGACPTCAGAAGGCCCCAAAGCCCTAACCTGTCCA 300  
 db 99985 TCCPAGGACCAAGGACACTGGAGACTCTGAGACTCTGAGAACCTAACCTGTCCA 100044  
 Qy 301 GCCAGAGATGGCTCTAGAGAGCTCTTCCAAAGCTTGTAGACAACCA 354  
 db 100045 GCCAGAGATGGCTCTAGAGAGCTCTTCCAAAGCTTGTAGACAACCA 1000998

RESULT 15

HSDJ977B1/c

LOCUS HSDJ977B1/c 145068 bp DNA Linear PRI 20-JUN-2001

DEFINITION Human DNA sequence from clone RP5-977B1 on chromosome 20 Contains ESTs, STSs, GSSs and three putative CG islands. Contains the 3' end of the gene for a novel protein tyrosine kinase, a heterogeneous nuclear ribonucleoprotein A3 pseudogene, the gene for three isoforms of a novel protein similar to putative RAB5-interacting protein, the TGF2 gene for TGF(beta)-induced transcription factor 2 with two isoforms, the MYRL2 gene for myosin regulatory light chain 2 (smooth muscle isoform), the 3' end of the gene KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4) with two isoforms and a novel gene, complete sequence.

ACCESSION AL05018

VERSION AL05018.13 GT:9581785

KEYWORDS HMG, CGP Island, heterogenous ribonucleoprotein, KIAA0964, myosin regulatory light chain, MYRL2, interacting protein, SH2 domain, TGF2, transcription factor, tyrosine kinase.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1. (bases 1 to 145068)

AUTHORS Lloyd.D.

TITLE Direct Submission

JOURNAL Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail: enquires: humquery@sanger.ac.uk Clone requests: ClonerRequest@sanger.ac.uk

COMMENT On Jul 28, 2000 this sequence replaced gi:5924017. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human

chromosome 20, constructed by the Sanger Centre Chromosome 20  
 Constructed by the Sanger Centre Chromosome 20  
 Further information can be found at  
[http://www.sanger.ac.uk/HGP/Chromosome\\_20](http://www.sanger.ac.uk/HGP/Chromosome_20)

This sequence is the entire insert of clone RP5-97731. The true left end of clone RP3-46029 is at 63218 in this sequence. The true right end of clone CTD-21819 is at 62218 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-977B1 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see  
<http://www.sanger.ac.uk/BLAST/home.htm>  
 VECTOR: pCYPAC2.

Location/Qualifiers

1..142068 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /clone="RP5-977B1"

568..875 /clone lib="RPCI-5"  
 /note="AlusX repeat: matches 1..311 of consensus"  
 1024..1580

repeat\_region 1047..1629 /note="match: GSS: Em: AQQ20498"  
 1631..1923 /note="match: GSS: Em: AQQ79162"  
 misc\_feature 2776..2842 /note="AlusSg repeat: matches 1..294 of consensus"  
 2776..2842 /note="TUN2 repeat: matches 2357..2423 of consensus"  
 3891..3949 /note="MIR repeat: matches 34..90 of consensus"  
 4456..4611 /note="MIR repeat: matches 67..231 of consensus"  
 5133..5549 /note="TUN2 repeat: matches 5924..6340 of consensus"  
 5550..5680 /note="FLAM\_C repeat: matches 1..132 of consensus"  
 5703..5890 /note="AlusSg repeat: matches 120..308 of consensus"  
 5923..6012 /note="TLMC\_D repeat: matches 5432..5528 of consensus"  
 6588..6889 /note="AlusP repeat: matches 1..303 of consensus"  
 34968..35059 /note="TUN2 repeat: matches 10394..10917..11329,  
 gene="KIAA0964"  
 /product="Alu977B1\_6..2 (KIAA0964 (ortholog of rat  
 PSD-95/SAP90-associated Protein 4))"  
 /note="match: cDNAs: Em: AB023181"  
 /evidence="not experimental"

repeat\_region 36569..36724..37531..39351) /gene="KIAA0964"  
 /product="Alu977B1\_6..1 (KIAA0964 (ortholog of rat  
 PSD-95/SAP90-associated Protein 4))"  
 /note="match: cDNAs: Em: AB000277 Em: AB000276  
 Em: AB000275 Em: AB023181 Em: AB000277 Em: AB000276  
 Em: U67137 Em: U67138 Em: U67987 Em: U67139 Em: AF0131778  
 match: ESTs: Em: A181867 Em: A162016 Em: AA196750  
 Em: AA154532 Em: A1747584 Em: A93116 Em: A88761  
 Em: AA154844 Em: A2272057 Em: A1202756 Em: A1491978  
 Em: A1519427 Em: A61165 Em: AAV6017 Em: A55820160  
 Em: A146940 Em: A190516 Em: T07912 Em: A1748854  
 Em: A1316293 Em: A1840413 Em: A1347117 Em: AA571014

FEATURES source

repeat\_region 3569..3624..37531..37749)

CDS

continues in dJ996C2 (AL445705) and bK2182L9 (AL390374)  
 match: proteins: Tr: Q9Y2H0"  
 /codon\_start=3  
 /evidence="not experimental"  
 /product="D1577B1\_6..2 (KIAA0964 (ortholog of rat  
 PSD-95/SAP90-associated protein 4))"  
 /protein\_id="ICAC16960\_1"  
 /db\_xref="GI:11323333..1"  
 /db\_xref="GQ:Q9Y2H0"  
 /db\_xref="SWISS-PROT:Q9Y2H0"  
 /translators="SSCLIVAYKTPPPVPPRTTSKPFISITVQSSTEASAQDTYLDSDQHKSFTMSGSGLNSSDSDSSRSPBPSUTRGVYAPAAPEAPEBPPRKHAALKSGQTLSQESHPAAPKRLSISGIGTBRTRNGSHLSEDGPKAIDVMAPSSESVPSMSRRDTSPTODANDSSCKSSENSLDPDTPHNSISDAGPQAPKTAQIRNLYSGDNDPALEASLLPDPWETSSSEPAQPAQCRACRFTKLIQDATERLEGNCQMDKE" KENNLSEETVSKLAYSQAQILMSQKFQPCQLONINDAANPPTAQDLAGFDLQLSIEDLSMKDDELYHAKRANSWOLVTPERKEKKPPKREPKSKPEAVSRDKA SDASDKORDAREKRLLAQKRAASVRONSAEADSTBEVYPAEATQ1" join(<743..7784..10306..10394..10917..11329..34968..35059,  
 36569..36724..37531..37749)

/note="KIAA0964"  
 /note="isoform 1"  
 continues in dJ996C2 (AL445705) and bK2182L9 (AL390374)  
 match: proteins: Tr: O4488 Tr: O14664 Tr: O1489 Tr: F97839  
 Tr: O14490 Tr: P97836 Tr: P97837 Tr: O54773 Tr: P97841  
 Tr: T78333"  
 /codon\_start=3  
 /evidence="not experimental"  
 /product="D1577B1\_6..1 (KIAA0964 (ortholog of rat  
 PSD-95/SAP90-associated protein 4))"  
 /protein\_id="ICAB570\_1"  
 /db\_xref="GI:16983730"  
 /db\_xref="GQ:Q9Y2H0"  
 /translators="SSCLIVAYKTPPPVPPRTTSKPFISITVQSSTEASAQDTYLDSDQHKSFTMSGSGLNSSDSDSSRSPBPSUTRGVYAPAAPEAPEBPPRKHAALKSGQTLSQESHPAAPKRLSISGIGTBRTRNGSHLSEDGPKAIDVMAPSSESVPSMSRRDTSPTODANDSSCKSSENSLDPDTPHNSISDAGPQAPKTAQIRNLYSGDNDPALEASLLPDPWETSSSEPAQPAQCRACRFTKLIQDATERLEGNCQMDKE" KENNLSEETVSKLAYSQAQILMSQKFQPCQLONINDAANPPTAQDLAGFDLQLSIEDLSMKDDELYHAKRANSWOLVTPERKEKKPPKREPKSKPEAVSRDKA SDASDKORDAREKRLLAQKRAASVRONSAEADSTBEVYPAEATQ1" DPALEASSLPPDWEETSSSEPAQPAQCRACRFTKLIQDATERLEGNCQMDKE" KENNLSEETVSKLAYSQAQILMSQKFQPCQLONINDAANPPTAQDLAGFDLQLSIEDLSMKDDELYHAKRANSWOLVTPERKEKKPPKREPKSKPEAVSRDKA SDASDKORDAREKRLLAQKRAASVRONSAEADSTBEVYPAEATQ1" join(<743..7784..10306..10394..10917..11329..34968..35059,  
 36569..36724..37531..37749)

/note="MIR repeat: matches 11..236 of consensus"  
 repeat\_region 11167..11178

repeat\_region 11178..11203

repeat\_region 11179..11203

repeat\_region 11204..11212

/note="L2 repeat: matches 2692..2708 of consensus"  
 /note="Alus repeat: matches 1..300 of consensus"  
 /note="L2 repeat: matches 2661..2692 of consensus"

```

repeat_region    12631. .12714   /note="HAI1 repeat: matches 1589. .1665 of consensus"
repeat_region    13077. .13367   /note="AluJ0 repeat: matches 3. .290 of consensus"
repeat_region    13440. .13742   /note="AluSX repeat: matches 1. .304 of consensus"
repeat_region    13932. .14197   /note="LIMB7 repeat: matches 5915. .6173 of consensus"
repeat_region    14198. .14512   /note="AluY repeat: matches 1. .307 of consensus"
repeat_region    14513. .14756   /note="LIMB7 repeat: matches 5650. .5915 of consensus"
repeat_region    14789. .15705   /note="LIMB4 repeat: matches 5162. .6185 of consensus"
repeat_region    15706. .16004

Query Match      12.8%; Score 152; DB 9; Length 145068;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy          778 AGGCTCTACTCTGTCAAGCCGCCTCAGCGGCCCTGCATCTGGACCGGATCACACA 837
Db          126094 AGGTCTACTCTGTCAAGCCGCCTGCATCTGGACCGGATCACACA 126035
Qy          838 CTACAGGATCACTGCCTTGACAATGSGCTGTAATCTACGGCCTGATCTGGACGATCACACA 897
Db          126034 CTACAGGATCACTGCCTTGACAATGSGCTGTAATCTACGGCCTGATCTGGACGATCACACA 125975
Qy          898 CTCACTCCAGGCCCTGGGACCATRACTCTG 929
Db          125974 CTCACTCCAGGCCCTGGGACCATRACTCTG 125943

```

Search completed: February 20, 2004, 07:24:47  
 Job time : 4588 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 01:41:55 (Without alignment) Search time 503 Seconds

Title: US-09-939-853A-74 Perfect score: 1183 Sequence: 1 agcttagctccaaggacc... . . . . . tctttggatgtgcctag 1183

Scoring table: OLIGO\_NTC GapP\_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseqn\_29Jan04:\*\*  
 1: Geneseqn1980s:\*\*  
 2: Geneseqn1990s:\*\*  
 3: Geneseqn000s:\*\*  
 4: Geneseqn001as:\*\*  
 5: Geneseqn2001as:\*\*  
 6: Geneseqn2002s:\*\*  
 7: Geneseqn2003as:\*\*  
 8: Geneseqn2003bs:\*\*  
 9: Geneseqn2003cs:\*\*  
 10: Geneseqn2004s:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

24	134	11.3	432	6	Abs02823 Human gen
25	134	11.3	448	4	Aai14520 Probe #44
	26	134	11.3	448	Abs0499 Human gen
	27	96	8.1	96	Aai22119 Probe #12
	28	96	8.1	96	Aba6198 Human fo
	29	96	8.1	96	Aai47414 Probe #16
	30	96	8.1	96	Aba4284 Human bre
	31	96	8.1	96	Aba34292 Probe #12
	32	96	8.1	96	Aak41374 Human bon
	33	96	8.1	96	Aak1540 Human bra
	34	96	8.1	96	Abs4966 Human liv
	35	96	8.1	96	Aai07818 Probe #78
	36	96	8.1	96	Abs15380 Human gen
	37	43	3.6	1348	Aal44087 Mouse mod
	38	30	2.5	30	Aad43983 Human leu
	39	26	2.2	26	Abk61505 Human NOV
	40	25	2.1	25	Aal44100 Human mod
	41	25	2.1	25	Aal4098 Human mod
	42	25	2.1	25	Aal44099 Human mod
	43	23	1.9	23	Aal44097 Human mod
	44	22	1.9	22	ABk61507 Human NOV
	45	22	1.9	22	Aax55277 Nucleotid

## ALIGNMENTS

RESULT 1				
ID	ABK61465	standard; cDNA;	1183 BP.	
XX				
AC	ABK61465;			
XX				
DT	18-JUN-2002	(first entry)		
XX				
DE	Human cDNA encoding protein NOV13.			
XX				
KW	Human; gene; ss: NOV; gene therapy; cardiomypathy; atherosclerosis; cell signal processing disorder; metabolic pathway modulation; disorder; diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;			
KW	uterus cancer; immune response; graft-versus-host disease; acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease; hypertension; congenital heart defects; multiple sclerosis; inflammation; Albright hereditary osteodystrophy.			
XX				
OS	Homo sapiens.			
XX				
PN	WC00216599-A2.			
XX				
PD	28-FEB-2002.			
XX				
XX	27-AUG-2001;	2001IW0-US026510.		
XX				
PR	25-AUG-2000;	2000US-D-0228191P.		
PR	08-FEB-2001;	2001US-D-0267310P.		
PR	20-FEB-2001;	2001US-D-0269861P.		
PR	20-MAR-2001;	2001US-D-0277337P.		
XX				
PA	(CURAGEN CORP.			
PA	(CORT-) COR THERAPEUTICS INC.			
PA	PA-280937/32.			
DR	P-PSDB; AAU1308.			
XX				
Burgess CB, Conley PB, Grosse WM, Hart M, Keltuda R, Shimkets RA;				
PI				
Sztek R, Szekeres ES, Tomlinson JE, Topper JN, Yang R;				
XX				
WPI;	2002-280937/32.			
DR	P-PSDB; AAU1308.			
XX				
New polypeptides for treating or preventing a disorder associated with them, in humans, e.g. cardiomypathy, atherosclerosis or cancers.				
XX				
Claim 1; Page 98; 263pp; English.				
PS				
CC	The invention relates to an isolated polypeptide (NOV) a mature form of			

## SUMMARIES

Result No.	Query	Score	Match Length	DB ID	Description
1	N_Geneseqn_29Jan04:**	1183	100.0	ABK61465	Abk61165 Human cDN
2		95.7	2567	6	Aad43980 Human SRC
3		62.1	786	6	Aal44089 Human mod
4		62.1	786	6	Abq74343 Human SRC
5		55.5	763	6	Abq98670 Human ORF
6		50.0	837	3	Aac77202 Human ORF
7		40.8	737	6	Aal44090 Mouse MAR
8		39.6	1413	6	Abq9934 Human cod
9		34.2	603	5	Abq99374 Human cod
10		33.9	2049	5	Aas74748 DNA encod
11		29.4	444	6	Abq99669 Human ORF
12		28.8	875	6	Abq99151 Human ORF
13		20.6	445	5	Aas74147 DNA encod
14		11.9	211	5	Aas70181 DNA encod
15		11.3	432	4	Aai12899 Probe #8
16		11.3	432	4	Aba4540 Human fo
17		11.3	432	4	Aai3236 Probe #29
18		11.3	432	4	Aba44128 Human bre
19		11.3	432	4	Aba24363 Probe #8
20		11.3	432	4	Aak28314 Human bon
21		11.3	432	4	Aak02872 Human bra
22		11.3	432	4	Ab22792 Human liv
23		11.3	432	4	Aai02797 Probe #27





Db	541	GATGAACTCTGGCTACTCAAGGAGCCCTGTGTCAGGGCTGCCCCCTCCCT	600
Qy	998	GCGAAGATAAACCTAACCTGTGACTTGCGAGGACACACTGAACGCTG	1057
Db	601	GCGAAGATAAACCTAACCTGTGACTTGCGAGGACACACTGAACGCTG	660
Db	1058	GCGAAGATAAACCTAACCTGTGACTTGCGAGGACACACTGAACGCTGAGGGT	1117
Qy	661	GCGAAGATAAACCTAACCTGTGACTTGCGAGGACACACTGAACGCTGAGCTG	720
Db	1118	CCTCGGAGTCCTCTGAGCTTACATAGCTGAAGGCTCTTTGATGAT	1177
Qy	721	CCTCGGAGTCCTCTGAGCTTACATAGCTGAAGGCTCTTTGATGAT	780
Qy	1178	GCCTAG 1183	
Db	781	GCCTAG 786	
RESULT 4			
	ABQ74343	standard; cDNA; 786 BP.	
ID	ABQ74343	standard; cDNA; 786 BP.	
XX			
AC	ABQ74343;		
XX			
DT	15-OCT-2002	{ first entry }	
DE	Human Src-like inhibitory molecule (SLIM) encoding cDNA.		
XX			
KW	Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;		
KW	SHAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;		
KW	modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;		
KW	acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;		
KW	chronic inflammatory disorder; autoimmune disorder; transplant rejection;		
KW	gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CD5	1. 786	
FT		/*tag= ^	
FT		/product= "SLIM"	
FT		/note= "Src-like inhibitory molecule"	
XX			
PN	WO200255707-A2.		
XX			
PD	18-JUL-2002.		
XX			
PF	10-JAN-2002; 2002WO-US000718.		
XX			
PR	10-JAN-2001; 2001US-0260953P.		
XX			
PA	(RIGE-) RIGEL PHARM INC.		
XX			
PI	Holland SJ, Mengenhall MK, Pardo J, Spencer C, Fu AC, Luo Y, N;		
PI	Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;		
XX			
DR	WPI; 2002-5/5432/61.		
DR	P-PSDB; ABP52187.		
XX			
PT	New src-like inhibitory molecule protein, useful for treating		
PT	immunodeficiency disorders and inflammatory disorders, comprises N-		
PT	terminal myristylation sequence, SH2 domain and/or SH3 domain.		
XX			
PS	Claim 3; Fig 2A; 91PP; English.		
XX			
CC	The present sequence encodes the human Src-like inhibitory molecule		
CC	(SLIM) protein [1]. The present invention describes a SLIM protein		
CC	comprising an N-terminal myristylation sequence, an N-terminal SH2		
CC	domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising		
CC	an N-terminal myristylation sequence and an N-terminal SH2 domain which		
CC	is unable to bind to Cbl. (1) has antiinflammatory, immunosuppressive and		
CC	immunosuppressive.		



allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive; ss.

*Homo sapiens*.

XX XX WO20058473-A2.

XX PN PD 05-OCT-2000.

XX XX 31-MAR-2000; 20000NO-US008621.

XX PR PR 31-MAR-1999; 990US-0127607P.  
PR 05-APR-1999;  
PR 05-APR-1999;  
PR 30-MAR-2000; 20000US-00540763.

(CUTRA-) CUTRAGEN CORP.

XX PA PI XX

Shimkets RA, Leach N;

XX DR WPI; 2000-602362/57.  
P-PSDB; AAB42993.

XX PT Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease.

XX PS Claim 5; Page 4692-4693; 5507pp; English.

XX AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The OREX sequences have activities such as: cytosolic; vulnerytic; antipsoratic; anticoronal; antiarachinic; immunosuppressant; immunostimulant; carcirant; thrombolytic; coagulant; vasoconstrictive; hypotensive; dermatological; immunosuppressive; antiflammatory; antibacterial; antiviral; antifungal; antineuritic; and antiaemic. These sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an OREX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vector. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular diseases diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; enhance coagulation; to inhibit thrombinosis; and as a contraceptive.

XX Semgene R37 RP: 176 A: 254 C: 245 G: 160 T: 0 U: 2 Other.

Y	679	GCTGTATGAGGGCTTGAAGGAAAGCAGGAACTGGTGTGTTACCTGGAAACC	738
P	243	GCTGTATGAGGGCTTGAAGGAAAGCAGGAACTGGTGTGTTACCTGGAAACC	302
Y	739	TGAGGGCCCTTCCTCACTCGGAGGCAACAGGAGGGCTCTACTCTCTGTCAGT	798
P	303	TGAGGGCCCTTCCTCACTCGGAGGCAACAGGAGGGCTCTACTCTCTGTCAGT	362
Y	799	CCGGCTAGGGCCCTGCATCCGGAGGCAACAGGAGGGCTCTACTGCTTGA	858
P	363	CCGGCTAGGGCCCTGCATCCGGAGGCAACAGGAGGGCTCTACTGCTTGA	422
Y	859	CAATGGTGGCTGTTACATTCACCGCGCTCACTTCCCCTCACTCAGGCCCTGTGAA	918
P	423	CAATGGTGGTTAATCTCACCGCGCTCACTTCCCCTCACTCAGGCCCTGTGAA	482
Y	919	CCATTACTCTGAGCTGGCATGACATCTGCTCTACTCAGGACCCCTGTGCTGAA	978
P	483	CCATTACTCTGAGCTGGCATGACATCTGCTCTACTCAGGACCCCTGTGCTGAA	542
Y	979	GGGGCTGGCCGGTCCCTGGAGGATAACCCCTAACCTGCTGTGAGAGAACCC	1038
P	543	GGGGCTGGCCGGTCCCTGGAGGATAACCCCTAACCTGCTGTGAGAGAACCC	602
Y	1039	ACTCACTGGAAGGAGCTGAAAGCTCCCTCTGGCAAGGGAGGA	1098
P	603	ACTCACTGGAAGGAGCTGAAAGCTCCCTCTGGCAAGGGAGGA	662
Y	1099	GTCCTCTCTGAGTGGGTCTGGGAGTCCTGACATAGCTGATGACGAA	1158
P	663	GTCCTCTCTGAGTGGGTCTGGGAGTCCTGACATAGCTGATGACGAA	722
Y	1159	GGCTGCTCTTGTGATGATGCTCTAG 1183	
P	723	GGCTGCTCTTGTGATGCTCTAG 747	

RESULTS 7  
LL44090      AAC44090      Standard: CDNA: 7337 BP.

A.D.44090.

卷之三

03-0CF-2002 (first entry)

Mouse MARS short isoform p

Mouse; gene; ss; gene therapy

MARS; tumour suppressor gene myeloid malignancy: acute

lumous expression; myeloproliferation;

Mus sp.

Key Concepts

CDS 1: 633

```
//tag=a  
/product=
```

W9399242452-A3

卷之三

30 - 2002.

26-NOV-2001; 2001WO-CA000166

27-NOV-2000; 2000CA-0232466

(HOSP-) HOSPITAL FOR SICK C

M. S. LANDER & J. C. LORENZO-MB.

卷之三

WPI; 2002-566564/60.  
P-PSDB; AAO15458.



in diagnostics as expressed sequence tags for identifying expressed genes. (II) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS6197-AAS9564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence	603 BP;	124 A;	189 C;	164 G;	126 T;	0 U;	0 Other;
Query Match	34.2%	Score 405;	DB 5;	Length 603			
Best Local Similarity	100.0%	Pred. No. 6	3e-161				
Matches	405;	Conservative	0;	Mismatches	0;	Indels	
779	GGCTCTACTCTCTGTCAGTCGCCCTAGCGCCCTGATCCTGGGAC						
199	GGCTCTACTCTCTGTCAGTCGCCCTGATCCTGGGAC						
839	TACAGGATTAATCTGCTTGAATGTGTCATGTACATGTCACCGGCC						
259	TACAGGATTAATCTGCTTGAATGGTGGCTGTCATGTCACCGGCC						
899	TOACTCGGGCCCTGGACCAATTACTTGACTGGGGATGACATCT						
319	TCACTCGCGCCCTGGACCAATTACTTGACTGGGGATGACATCT						
959	AAGGACGCCCTGTCTGGAGGGCTGCCGCTGGCCGAGGATA						
379	AAGGGACGCCCTGTCTGGAGGGCTGCCGCTGGCCGAGGATA						
1019	GTTGACTGTGAGAGGACACCACTGAACTGAAAGGTGACAGCTCCC						
439	GTGACTGTGAGAGGACACCACTGAAAGGTGACAGCTCCC						
1079	GAAGCTGCATAGGGAGGAGTCTTCTCAGTAGGGCTCTGGGAGT						
499	GAAGCTGCACAGGGAGGAGTCTTCTCAGTAGGGCTCTGGGAGT						
1139	TACATGCGCTGAATGAGGGCTCTCTTGGATGATGCCCTAG 118						
559	TACATGCGCTGAATGAGGGCTCTCTTGGATGATGCCCTAG 503						

SJSU-IT 10  
S7750  
AAS74750 standard; cDNA; 2049  
AAS74750;  
13-FEB-2002 (first entry)  
DNA encoding novel human diag  
Human: chromosome mapping: ce

30-MAR-2001: 2001IWO-US0008631.  
11-OCT-2001.  
Homo sapiens.  
WO200175067-A2.

31-MAR-2000; 2000US-00540217.  
23-AUG-2000; 2000US-00649167.  
  
(HYSE-) HYSEQ INC.  
  
Dermanac RT, Liu C, Tang YT;  
WPI, 2001-633362/73.  
P-PSDB; ABG10563.  
  
New isolated polynucleotide as  
diagnostics, forensics, gene  
responsible for Genetic disorder  
biodiversity.

Claim 1; SEQ ID NO 10554; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences, (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS61197-AAS94164 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences)

Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;





OS Homo sapiens.  
 XX WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000015-00540217.  
 XX DRMANAC RT, Liu C, Tang YT;  
 XX PI PI  
 XX DR WPI: 2001-639162/73.  
 XX P-PSDB; ABG05594.  
 XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.  
 XX PS Claim 1; SEQ ID NO 5985; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a supplement in tissue, as molecular weight markers and as a food supplement (II) and its binding partners are useful in medical imaging of sites expressing (II); (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-DAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 211 BP; 50 A; 51 C; 72 G; 38 T; 0 U; 0 Other;  
 XX Query Match 11.9%; Score 141; DB 5; Length 211;  
 Best Local Similarity 9.5%; Pred. No. 3.5e-56;  
 Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 YY 588 AGGAGCGAGCTGGTGGACGCTGTGAACTCTCAGGCCAGAGCTATAACATCCCCA 647  
 Db 20 AGGATGAGCTGGTGGAGCTGGTGTGAACTCTCAGGCCAGAGCTATAACATCCCCA 79  
 YY 648 GCGTCACGGGGCAAAGTCTCCCATGGTGTGAAAGGGCTTGAGGCCCTGAGGGGGAAAG 707  
 Db 80 CGTCACGGGGCAAAGTCTCCCATGGTGTGAAAGGGCTTGAGGCCCTGAGGGGGAAAG 139  
 YY 708 CAGAGGAAGCTGTGTGAACTCTGGAAACCTGAGGGCTTGAGGCCCTGAGGGGGAAAG 767  
 Db 140 CAGAGGAAGCTGTGTGAACTCTGGAAACCTGAGGGCTTGAGGCCCTGAGGGGGAAAG 199  
 YY 768 AGACCAAGGAG 779  
 Db 200 AGACCAAGGAG 211

XX AAI12879;  
 XX D12-OCT-2001 (first entry)  
 DE Probe #2812 for gene expression analysis in human cervical cell sample.  
 XX Probe; human; microarray; gene expression; cervical epithelial cell;  
 XX cervical cancer; ss.  
 OS Homo sapiens.  
 XX PN WO200157278-A2.  
 XX DR 09-AUG-2001.  
 XX PPF 30-JAN-2001; 2001WO-US000670.  
 XX PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-036359P.  
 PR 04-OCT-2000; 2000GB-0004263.

PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PA XX DR WPI: 2001-488901/53.  
 PR Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human cervical epithelial cells.  
 PA XX Claim 25; SEQ ID NO 2812; 487pp; English.  
 PR XX The present invention relates to human single exon nucleic acid probes (SENP). The present sequence is one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC XX Sequence 432 BP; 109 A; 115 C; 97 G; 97 T; 0 U; 0 Other;  
 CC Query Match 11.3%; Score 134; DB 4; Length 432;  
 CC Best Local Similarity 100.0%; Pred. No. 6.8e-53;  
 CC Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC YY 929 GAGCTGGATGAGCTGGTGTGAACTCTGGCTACTTAAGGAGCCCTGTCCTGAGGGCTGC 988  
 CC Db 270 GAGCTGGATGAGCTGGTGTGAACTCTGGCTACTTAAGGAGCCCTGTCCTGAGGGCTGC 329  
 CC YY 989 CGCTCCCTGCCAGGATAACCCCTACCTGTGACTGTGTCAGGGACACACTCAACTTG 1048  
 CC Db 330 CGCTCCCTGCCAGGATAACCCCTACCTGTGACTGTGTCAGGGACACACTCAACTTG 389  
 CC YY 1049 AAAGAGCTGAGCAG 1062  
 CC Db 390 AAAGAGTTGGAGCAG 403

Search completed: February 20, 2004, 06:07:58  
 Job time : 512 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run cn: February 20, 2004, 05:08:33 ; Search time 2929 Seconds

(without alignments)

12061.099 Million cell updates/sec

Title: US-09-939-853a-74

Perfect score: 1183

Sequence: 1 agcttagactccaaaggaccc.....tctctttggatgtgcctag 1183

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:\*

1: em\_estba:\*

2: em\_estbum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gib\_est:\*

10: gib\_est2:\*

11: gib\_hmc:\*

12: gib\_est3:\*

13: gib\_est4:\*

14: gib\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss hum:\*

18: em\_gss inv:\*

19: em\_gss\_pn:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

em\_gss\_pig:\*

27: gib\_gbs1:\*

28: gib\_gbs2:\*

29: gib\_gbs2:\*

30: gib\_gbs2:\*

31: gib\_gbs2:\*

32: gib\_gbs2:\*

33: gib\_gbs2:\*

34: gib\_gbs2:\*

35: gib\_gbs2:\*

36: gib\_gbs2:\*

37: gib\_gbs2:\*

38: gib\_gbs2:\*

39: gib\_gbs2:\*

40: gib\_gbs2:\*

41: gib\_gbs2:\*

42: gib\_gbs2:\*

43: gib\_gbs2:\*

44: gib\_gbs2:\*

45: gib\_gbs2:\*

46: gib\_gbs2:\*

47: gib\_gbs2:\*

48: gib\_gbs2:\*

49: gib\_gbs2:\*

50: gib\_gbs2:\*

51: gib\_gbs2:\*

52: gib\_gbs2:\*

53: gib\_gbs2:\*

54: gib\_gbs2:\*

55: gib\_gbs2:\*

56: gib\_gbs2:\*

57: gib\_gbs2:\*

58: gib\_gbs2:\*

59: gib\_gbs2:\*

60: gib\_gbs2:\*

61: gib\_gbs2:\*

62: gib\_gbs2:\*

63: gib\_gbs2:\*

64: gib\_gbs2:\*

65: gib\_gbs2:\*

66: gib\_gbs2:\*

67: gib\_gbs2:\*

68: gib\_gbs2:\*

69: gib\_gbs2:\*

70: gib\_gbs2:\*

71: gib\_gbs2:\*

72: gib\_gbs2:\*

73: gib\_gbs2:\*

74: gib\_gbs2:\*

75: gib\_gbs2:\*

76: gib\_gbs2:\*

77: gib\_gbs2:\*

78: gib\_gbs2:\*

79: gib\_gbs2:\*

80: gib\_gbs2:\*

81: gib\_gbs2:\*

82: gib\_gbs2:\*

83: gib\_gbs2:\*

84: gib\_gbs2:\*

85: gib\_gbs2:\*

86: gib\_gbs2:\*

87: gib\_gbs2:\*

88: gib\_gbs2:\*

89: gib\_gbs2:\*

90: gib\_gbs2:\*

91: gib\_gbs2:\*

92: gib\_gbs2:\*

93: gib\_gbs2:\*

94: gib\_gbs2:\*

95: gib\_gbs2:\*

96: gib\_gbs2:\*

97: gib\_gbs2:\*

98: gib\_gbs2:\*

99: gib\_gbs2:\*

100: gib\_gbs2:\*

101: gib\_gbs2:\*

102: gib\_gbs2:\*

103: gib\_gbs2:\*

104: gib\_gbs2:\*

105: gib\_gbs2:\*

106: gib\_gbs2:\*

107: gib\_gbs2:\*

108: gib\_gbs2:\*

109: gib\_gbs2:\*

110: gib\_gbs2:\*

111: gib\_gbs2:\*

112: gib\_gbs2:\*

113: gib\_gbs2:\*

114: gib\_gbs2:\*

115: gib\_gbs2:\*

116: gib\_gbs2:\*

117: gib\_gbs2:\*

118: gib\_gbs2:\*

119: gib\_gbs2:\*

120: gib\_gbs2:\*

121: gib\_gbs2:\*

122: gib\_gbs2:\*

123: gib\_gbs2:\*

124: gib\_gbs2:\*

125: gib\_gbs2:\*

126: gib\_gbs2:\*

127: gib\_gbs2:\*

128: gib\_gbs2:\*

129: gib\_gbs2:\*

130: gib\_gbs2:\*

131: gib\_gbs2:\*

132: gib\_gbs2:\*

133: gib\_gbs2:\*

134: gib\_gbs2:\*

135: gib\_gbs2:\*

136: gib\_gbs2:\*

137: gib\_gbs2:\*

138: gib\_gbs2:\*

139: gib\_gbs2:\*

140: gib\_gbs2:\*

141: gib\_gbs2:\*

142: gib\_gbs2:\*

143: gib\_gbs2:\*

144: gib\_gbs2:\*

145: gib\_gbs2:\*

146: gib\_gbs2:\*

147: gib\_gbs2:\*

148: gib\_gbs2:\*

149: gib\_gbs2:\*

150: gib\_gbs2:\*

151: gib\_gbs2:\*

152: gib\_gbs2:\*

153: gib\_gbs2:\*

154: gib\_gbs2:\*

155: gib\_gbs2:\*

156: gib\_gbs2:\*

157: gib\_gbs2:\*

158: gib\_gbs2:\*

159: gib\_gbs2:\*

160: gib\_gbs2:\*

161: gib\_gbs2:\*

162: gib\_gbs2:\*

163: gib\_gbs2:\*

164: gib\_gbs2:\*

165: gib\_gbs2:\*

166: gib\_gbs2:\*

167: gib\_gbs2:\*

168: gib\_gbs2:\*

169: gib\_gbs2:\*

170: gib\_gbs2:\*

171: gib\_gbs2:\*

172: gib\_gbs2:\*

173: gib\_gbs2:\*

174: gib\_gbs2:\*

175: gib\_gbs2:\*

176: gib\_gbs2:\*

177: gib\_gbs2:\*

178: gib\_gbs2:\*

179: gib\_gbs2:\*

180: gib\_gbs2:\*

181: gib\_gbs2:\*

182: gib\_gbs2:\*

183: gib\_gbs2:\*

184: gib\_gbs2:\*

185: gib\_gbs2:\*

186: gib\_gbs2:\*

187: gib\_gbs2:\*

188: gib\_gbs2:\*

189: gib\_gbs2:\*

190: gib\_gbs2:\*

191: gib\_gbs2:\*

192: gib\_gbs2:\*

193: gib\_gbs2:\*

194: gib\_gbs2:\*

195: gib\_gbs2:\*

196: gib\_gbs2:\*

197: gib\_gbs2:\*

198: gib\_gbs2:\*

199: gib\_gbs2:\*

200: gib\_gbs2:\*

201: gib\_gbs2:\*

202: gib\_gbs2:\*

203: gib\_gbs2:\*

204: gib\_gbs2:\*

205: gib\_gbs2:\*

206: gib\_gbs2:\*

207: gib\_gbs2:\*

208: gib\_gbs2:\*

209: gib\_gbs2:\*

210: gib\_gbs2:\*

211: gib\_gbs2:\*

212: gib\_gbs2:\*

213: gib\_gbs2:\*

214: gib\_gbs2:\*

215: gib\_gbs2:\*

216: gib\_gbs2:\*

217: gib\_gbs2:\*

218: gib\_gbs2:\*

219: gib\_gbs2:\*

220: gib\_gbs2:\*

221: gib\_gbs2:\*

222: gib\_gbs2:\*

223: gib\_gbs2:\*

224: gib\_gbs2:\*

225: gib\_gbs2:\*

226: gib\_gbs2:\*

227: gib\_gbs2:\*

228: gib\_gbs2:\*

229: gib\_gbs2:\*

230: gib\_gbs2:\*

231: gib\_gbs2:\*

232: gib\_gbs2:\*

233: gib\_gbs2:\*

234: gib\_gbs2:\*

235: gib\_gbs2:\*

236: gib\_gbs2:\*

237: gib\_gbs2:\*

FEATURES	source	Query Match	Score	Match	Length
http://fulllength.invitrogen.com/ Invitrogen.com/ Faraday Avenue Genoscope Sequence ID : CSODE005AF12QPL.					1600
Location/Qualifiers					
1..1201					
/organism="Homo sapiens"					
/mol type="mRNA"					
/db_Xref="taxon: 9606"					
/clone="CSODE005YK23"					
/tissue="Tib="Homo sapiens PLACENTA"					
/note="Vector: PCMVSPORT 6; 1st strand cDNA was primed with a No.1-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRI sites of the PCMVSPORT 6 vector. Library was not normalized."					
ORIGIN					
		Query Local Similarity	62.1%	Score	735;
		Matches	835;	Pred. No.	0;
		Conservative	0;	Mismatches	2;
		Indels	0;	Gaps	0;
DY	b	4 TAGAGCTCCAAGGACCCACCGCTGTGAGCTCAAGAGCTCAAAGGGGCCCTGGGCCT	63		
	b	65 TAGAGCTCCAAGGACCCACCGCTGTGAGCTCAAGAGCTCAAAGGGGCCCTGGGCCT	124		
	b	64 TCCCTCCCTGCTGTGCTGGGTTCCTCCAGTCAGATCCTCAAAGGAGCAT	123		
	b	125 TCCCTCCCTGCTGTGCTGGGTTCCTCCAGTCAGATCCTCAAAGGAGCAT	184		
	b	124 GGGGAGCTGATCCATCCCTGTTACAATTGCTGACTCGAGAAGATGCTGAGTACC	183		
	b	185 GGGGAGCTGATCCATCCCTGCTGAGATCCTCCAGGTGAGAGAGTCTGGGTGTC	244		
	b	184 CAAACCAACACCTAGCTCCTGAGCTGAGATCCTCCAGGTGAGAGAGTCTGGGTGTC	243		
	b	245 CAAACCAACACCTAGCTCCTGAGATGCTCCCTGAGATCTCCAGGTGAGAGAGTCTGGGTGTC	304		
	b	244 TAGGACCAAGGACACTGGCACACTTGAGAGCTTCTGAGAGCTAACCTGTCAGGCC	303		
	b	305 TAGGACCAAGGACACTGGCACACTTGAGAGCTTCTGAGAGCTAACCTGTCAGGCC	364		
	b	304 AGAGCATGGCTCTGAGAGCTGTTCCAAAGCTTGTGAGAACCAATTTCCTC	363		
	b	365 AGAGCATGGCTCTGAGAGCTGTTCCAAAGCTTGTGAGAACCAATTTCCTC	424		
	b	364 GATGAAGTGTCTGAGTGTGCTGAGTGGACAATGGAGTCTGAGTGGAA	423		
	b	425 GATGAAGTGTCTGAGTGTGCTGAGTGGACAATGGAGTCTGAGTGGAA	484		
	b	424 ATCTCTGCGCAAGGCCAGGCTGAGTGTGAGTGTGAGTGTGAGTGGAA	483		
	b	485 ATCTCTGCGCAAGGCCAGGCTGAGTGTGAGTGTGAGTGTGAGTGGAA	544		
	b	484 AGCAAGGAGAAAGCCAAGGCCATGCCCTGGCCAGTTCCCGAACGGCCGG	543		
	b	545 AGCAAGGAGAAAGCCAAGGCCATGCCCTGGCCAGTTCCCGAACGGCCGG	604		
	b	544 CGAGCTGCTGCTGAGTGTGAGTGTGAGTGTGAGTGGAGACTGGTG	603		
	b	605 CGAGCTGCTGCTGAGTGTGAGTGTGAGTGTGAGTGGAGACTGGTG	663		
	b	604 GACGGCTCTGCTGAGTGTGAGTGTGAGTGTGAGTGGAGACTGGTG	664		
	b	665 GACGGCTCTGCTGAGTGTGAGTGTGAGTGTGAGTGGAGACTGGTG	724		
	b	664 AGTCCTCCATGGCTGGCTGTATGAGGGCTCTGAGCTGGAGAAACTCTGTT	723		
	b	725 AGTCCTCCATGGCTGGCTGTATGAGGGCTCTGAGCTGGAGAAACTCTGTT	784		
	b	724 GTTACCTGGAAACCTCTGAGGGCTCTGAGCTGGAGCTGGAGCTGGCTC	783		
	b	785 GTTACCTGGAAACCTCTGAGGGCTCTGAGCTGGAGCTGGAGCTGGCTC	844		

Qy	784	TATCTCTGTCAGTCGGCTCAIGCGCCCTGATCTGGACGGATCAGACACTA	840
Db	845	TATCTCTGTCAGTCGGCTCAIGCGCCCTGATCTGGACGGATCAGACACTA	901
<b>RESULT 2</b>			
BQ052308	BQ052308	1002 bp	mRNA linear EST 29-MAR-2002
LOCUS	AGENCOURT 6868571	NIH_MGC_106	Homo sapiens cDNA clone IMAGE:5933542
DEFINITION	5' - mRNA Sequence.		
ACCESSION	BQ052308		
VERSION	BQ052308.1	GI:19811648	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
MATERIAL	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 - (bases 1 to 1002)		
AUTHORS	NIH-MGC	http://mgc.ncbi.nlm.nih.gov/	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgbps@mail.nih.gov Tissue Procurement: Dr. Daniel McVicar, DBS/NCI cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution by: Agencourt Bioscience Corporation Information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov		
Plate:	LICM2118	row: d column: 23	
High quality sequence stop: 670.			
Location/Qualifiers			
FEATURES	1..1002		
Source	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="Exon.9606" /clone="IMAGE:5933542" /tissue type="natural killer cells, cell line" /lab host="IDH10B (phage-resistant)" /clone lib="NIH_MGC_106" /note="Organ: blood; Vector: pORT7; Site_1: XbaI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoR1/XbaI sites using the following 5' adaptor: GGCAAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
<b>ORIGIN</b>			
Query Match	54.7%	Score 647; DB 12; Length 1002;	
Best Local Similarity	99.7%	Pred. No. 0;	
Matches	747; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
Qy	333	CCAGGCCCTTGATGACAACTTCCCTGGATGATGTTGCTCTGCTCTGCTGAG 392	
Db	31	CCAGGCCCTTGATGACAACTTCCCTGGATGATGTTGCTCTGCTCTGCTGAG 90	
Qy	393	GAACTATGGAAAGTCTGCCAGCAGAAATACTCTGCAAGCCAAGCTTGAGTCTCT 452	
Db	91	GAACATGGAAAGTCTGCCAGCAGAAATACTCTGCAAGCCAAGCTTGAGTCTCT 150	
Qy	453	CCTGGCGCATTTCCGGAGGTGGCCGGCCGAGCTGCTGAGACTCGGGAGCCAT 572	
Db	151	CCTGGCGCATTTCCGGAGGTGGCCGGCCAGCTGCTGAGACTCGGGAGCCAT 210	
Qy	513	CCTGGCGCATTTCCGGAGGTGGCCGGCCGAGCTGCTGAGACTCGGGAGCCAT 572	
Db	211	CCTGGCGCATTTCCGGAGGTGGCCGGCCAGCTGCTGAGACTCGGGAGCCAT 270	
Qy	573	TGACCATGCTCTGAGATGGGACTGTGGCTGAGACTCTGGCTGAATCTGGCAAG 632	

Db 271 TGGACCATCGTCCTGAGGATGGAGACTGGTGGACGGTGTGAGTCAGGGAGAG 330  
 Qy 633 AGTAACATCCAGCGTCACTGGCAAAAGTCTCCATGGTGGCTGTAGGGGCC 692  
 Db 331 AGTATAACATCCCCAGGTCACTGGCAAAAGTCCATGGTGGCTGTAGGGCC 390  
 Qy 693 TGAGCAGGGAAAGAACAGGAATCTGGAAACCTGGAGGGCTTCC 752  
 Db 391 TGAGCAGGGAAAGAACAGGAATCTGGAAACCTGGAGGGCTTCC 450  
 Qy 753 TCATCGGGAGACGACCCATTCTACTCTGCAATGGCTCACGCC 812  
 Db 451 TCATCGGGAGACGACCCATTCTACTCTGCAATGGCTCACGCC 510  
 Qy 813 CTGCACTCTGGACCGGATCAAGACAATGGCTGTGT 872  
 Db 511 CTGCACTCTGGACCGGATCAAGACAATGGCTGTGT 570  
 Qy 873 ACATTCACCGGCCTCACCTCCCTCACTCCAGGCGCTGGGACCAATTAACTTGAGC 932  
 Db 571 ACATTCACCGGCCTCACCTCCCTCACTCCAGGCGCTGGGACCAATTAACTTGAGC 630  
 Qy 933 TGGCTGATGACATCTGCGCCTCACCTCCCTCACTCCAGGCGCTGGGACCCAC 992  
 Db 631 TGGCTGATGACATCTGCGCCTCACCTCCCTCACTCCAGGCGCTGGGACCCAC 690  
 Qy 993 TCCCCTGGCAAGGATAAACCCCTACCTCTGACTGTGGAGGACCAACTCACTTGAAAG 1052  
 Db 691 TCCCCTGGCAAGGATAAACCCCTACCTCTGACTGTGGAGGACCAACTCACTTGAAAG 750  
 Qy 1053 AGCTGGAGACGCTCCCTCTGACTGTGGAGGACCAACTCACTTGAAAG 1081  
 Db 751 AGCTGGAGACGCTCCCTCTGACTGTGGAGGACCAACTCACTTGAAAG 779

RESULT 3  
 BQ054265 LOCUS BQ054265 986 bp mRNA linear EST 29-MAR-2002  
 DEFINITION AGENCOURT\_6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
 ACCSSION BQ054265 5', mRNA sequence.  
 VERSION BQ054265.1 GI:19813605  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Mammalia; Buthidae; Chordata; Craniata; Vertebrata; Buteleostomi;  
 REFERENCE 1 (bases 1 to 986)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgsbps@mail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L1Cn2125 row: i column: 12  
 High quality sequence stop: 515.

FEATURES Location/Qualifiers  
 1. .986  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="Taxon:9606"  
 /clone IMAGE:5936339  
 /tissue type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XbaI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned"

SOURCE

RESULT 4  
 BQ054281 LOCUS BQ054281 1020 bp mRNA linear EST 29-MAR-2002  
 DEFINITION AGENCOURT\_6830234 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936362  
 5', mRNA sequence.  
 ACCESSION BQ054281  
 VERSION BQ054281.1 GI:19813621  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1020)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

COMMENT	Contact: Robert Straussberg, Ph.D. Tissue Procurement: Dr. Daniel McVicar, DBS/NCI DNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Clone distribution by: Agencourt Bioscience Corporation found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>	Qy 791 CTGTCTAGTCGGCCTCAGGCGCCCTGATCCTGGACGGATCAGAACATACAGGATCCAC 850 Db 605 CTGTCTAGTCGGCCTCAGGCGCCCTGATCCTGGACGGATCAGAACATACAGGATCCAC 664
FEATURES	Plate: LLCM2125 row: j column: 11 High quality sequence stop: 556. Location/Qualifiers 1..1020 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:3606" /clone:"IMAGE:5936362" /tissue_type="natural killer cells, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC_106" /ncce="Organ: b10cd; vector: pOTB7; site1: XbaI; site2: EcoRI; cDNA made by ECO RI cloning sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript III RT (Life Technologies). Note: this is a NIH MGC Library."	Qy 851 TGCTCTAGTCGGCCTCAAGGCGCCAAAGCCCTAACCTGTCAGCCAGAGCAT 910 Db 665 TGCTCTAGTCGGCCTCAAGGCGCCAAACTCTGTCAGCCAGAGCAT 724
ORIGIN	Query Match 48 0%; Score 568; DB 12; Length 1020; Best Local Similarity 99.6%; Pred. No. 1.5e-276 ; Matches 718; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	Qy 191 ACACCTAGCTCTCCCTGAATCTCCGGCTTAGAACCC 250 Db 5 ACACCTAGCTCTCCCTGAATCTCCGGCTTAGAACCTGGCTTAGAACCC 64 Qy 251 AAGGACACTGGCAGACTTCAGAAGGGCCCAAAGCCCTAACCTGTCAGCCAGAGCAT 310 Db 65 AAGGACACTGGCAGACTTCAGAAGGGCCCAAAGCCCTAACCTGTCAGCCAGAGCAT 124 Db 125 GGGTCTAGZAGAGGTGTCTCCAGCTTGTCTGGTCACTGGAGATG 184 Qy 311 GGTCTAGZAGAGGTGTCTCCAGCTTGTCTGGTCACTGGAGATG 370 Db 125 GGGTCTAGZAGAGGTGTCTCCAGCTTGTCTGGTCACTGGAGATG 184 Qy 371 TGCTCTAGTCGGCCTCAAGGAAATCTGGCAAGAACATCTCTG 430 Db 185 TGCTCTAGTCGGCCTCAAGGAAATCTGGCAAGAACATCTCTG 244 Qy 431 CCAAGGCCAAGGTGAGTTCTGTCACAGCCAGGACCTGTGACCATGGAGCAG 490 Db 245 CCAAGGCCAAGGTGAGTTCTGTCACAGCCAGGACCTGTGACCATGGAGCAG 304 Qy 491 AGAAAGCAAGGCCACAGCCTGGCCAGTTCCCGCAGGTCGAGCTG 550 Db 305 AGAAAGCAAGGCCACAGCCTGGCCAGTTCCCGCAGGTCGAGCTG 364 Qy 551 TGGCTGAGACTGGGGGCAATTGACATCTGAGGATGGAGCTGGAGCTG 610 Db 365 TGGCTGAGACTGGGGGCAATTGACATCTGAGGATGGAGCTGGAGCTG 424 Qy 611 CGTCTGAATGTCAGGCGAGGATTAACATCCCAGGTCACGTGGCAAGTCTCC 670 Db 425 CGTCTGAATGTCAGGCGAGGATTAACATCCCAGGTCACGTGGCAAGTCTCC 484 Qy 671 CATGGGGCTGTTAGGGCTGAGAGGAGAANGCAAGGAGACTGCTGTTACCT 730 Db 485 CATGGGGCTGTTAGGGCTGAGAGGAGAANGCAAGGAGACTGCTGTTACCT 544 Qy 731 GGGAAACCTGGAGGGCTTCTCTCATGGAGAGCAACGGAGGGCTTACTCT 790 Db 545 GNGAACCCCTGGAGGGCTTCTCTCATGGAGAGCAACGGAGGGCTTACTCT 604
source	RESULT 5 BQ052468 LOCUS BQ052468 1069 bp mRNA linear EST 29-MAR-2002 DEFINITION AGENCOURT 6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772 KEYWORDS 5', mRNA sequence. ACCESSION BQ052468 VERSION BQ052468.1 GI:19811808 SOURCE Homo sapiens (human) ORGANISM Homo sapiens KEYWORDS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE NIH_MGC http://msc.nci.nih.gov/ AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC) TITLE NIH_MGC Unpublished (1999) JOURNAL Contact: Robert Straussberg, Ph.D. COMMENT Unpublished (1999) Email: cgappa@nigms.nih.gov Tissue Procurement: Dr. Daniel McVicar, DBS/NCI cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) DNA Segmentation by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2118 row: n column: 13 High quality sequence stop: 681. Location/Qualifiers 1..1069 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone:"IMAGE:5933772" /tissue_type="natural killer cells, cell line" /lab_host="DH10B (phage resistant)" /note="Organ: blood; Vector: pOTB7; Site1: XbaI; Site2: EcoRI; cDNA made by oligo-dT Priming. Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and SuperScript II RT (Life Technologies). Note: this is a NIH_MGC Library."	Qy 310 TGGCTGAGCTGGCTCTGGTCACTGGAGAATCTGGCTGGATG 369 Db 1 TGGCTGAGCTGGCTCTGGTCACTGGAGAATCTGGCTGGATG 60 Qy 370 GTGGTTCTGAGCTGGCTCTGGTCACTGGAGAATCTGGCTGGATG 429 Db 61 GTGGTTCTGAGCTGGCTCTGGAGAACATGGCTGGAGAATCTGGCTGGATG 120 Qy 430 GCCAAGCCCCAAGGTTGAGTGGCTCTGGCTGGAGGACTCTGGTACCACTGGAGAAGGAGA 489

10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

Db	121	GCCAAAGCCAAAGCTGAGTTCCTGCAAGCCTGTGACATGAAAGCGA	180
Qy	490	SAGAACAAAGCCACAGCTGGCCCTGGCCAGTTTCCGGCAGGTGCGGCAGCT	549
Db	181	SAGAACAAAGCCACAGCCGGCCCTGGCAAGTTCCTGGCGAGCT	240
Qy	550	GTCGCTGAGATCGGGAGCATTGACCATGTCAGATGGAGACTGGTGACGCT	649
Db	241	GTCGCTGAGATCGGGAGCATTGACCATGTCAGATGGAGACTGGTGACGCT	300
Qy	610	GTCGCTGAGATCGGGAGCATTGACCATGTCAGATGGAGACTGGTGACGCT	669
Db	301	GTCGCTGAGATCGGGAGCATTGACCATGTCAGATGGAGACTGGTGACGCT	360
Qy	670	CCATGGCTGAACTCTAGGAGAGATACTCCAGGTCACTGGCAAGTCTC	729
Db	361	CCATGGCTGAACTCTAGGAGAGATACTCCAGGTCACTGGCAAGTCTC	420
Qy	730	TGGAAACCTCTAGGGCCTTCCTCATCCGGAGAGGCTCTTAATCTC	789
Db	421	TGGAAACCTCTAGGGCCTTCCTCATCCGGAGAGGCTCTTAATCTC	480
Qy	790	TCTGTCACTGGCTCTAGGCCCTGATCTGGACCGGATCACAGATCCA	849
Db	481	TCTGTCACTGGCTCTAGGCCCTGATCTGGACCGGATCACAGATCCA	540
Qy	850	CTGCCCTGACATGGCTGGCTGCTACATCTCACCCCTACCTCCACAGGC	909
Db	541	CTGCCCTGACATGGCTGGCTGCTACATCTCACCCCTACCTCCACAGGC	600
Qy	910	CCTGG 914	
Db	601	CCTGG 605	

ORIGIN

Query	Match	Score	Length
Best	Local	Similarity	616;
Matches	509;	Conservative	
Qy	9	CTCCAAAGACCCACCGCTGCTCTGTGAGCTCAAAGGGCCCTGGCCCT	68
Db	58	CTCCAAAGACCCACCGCTGCTCTGTGAGCTCAAAGGGCCCTGGCCCT	117
Qy	69	CCCTGGCTGGCTGGCTGGTGGAGGGTCCCAGTCAGATCCCTAAGGAGC	128
Db	118	CCCTGGCTGGCTGGCTGGTGGAGGGTCCCAGTCAGATCCCTAAGGAGC	177
Qy	129	AGTGTATCATCCCTGTTGTAACAATCTGCTACTGAGAGATGTGAGTAC	188
Db	178	AGTGTATCATCCCTGTTGTAACAATCTGCTACTGAGAGATGTGAGTAC	237
Qy	189	CAACACCTAGCCCTCCCTGAGATCTCCAGGTGAGAGTCTGGGTGCTTAGA	248
Db	238	CAACACCTAGCCCTCCCTGAGATCTCCAGGTGAGAGTCTGGGTGCTTAGA	297
Qy	249	CCAAGGACACTGGCAGACTCCAGAAGGCCCAAGGCCCTAACCTGTCAGGAGC	308
Db	298	CCAAGGACACTGGCAGACTCCAGAAGGCCCAAGGCCCTAACCTGTCAGGAGC	357
Qy	309	ATCGCTCTCACAGAGAGCTGCTCTCCAGAAGACAACTTCCCTTCGATGA	368
Db	358	ATCGCTCTCACAGAGAGCTGCTCTCCAGAAGACAACTTCCCTTCGATGA	417
Qy	369	TGNGCTTGTGAGTGGAAACATGGGAGCTGGCCAGAGAAATACTTC	428
Db	418	TGNGCTTGTGAGTGGAAACATGGGAGCTGGCCAGAGAAATACTTC	477
Qy	429	TGCAAGGCCAACCTGGATTCCTGGCTGGAGAACATGGAACTTGCCAGA	488
Db	478	TGCAAGGCCAACCTGGATTCCTGGCTGGAGAACATGGAACTTGCCAGA	537
Qy	489	AGAAAGGAAGGCCAACGGCTGGCCCTGG 518	
Db	538	AGAAAGGAAGGCCAACGGCTGGCCCTGG 567	

RESULT 7

LOCUS	BX383605	616 bp mRNA	EST 08-MAY-2003
DEFINITION	BX383606 Homo sapiens T CELLS (JURKAT CELL LINE)	COT 10-NORMALIZED	
ACCESSION	BX383606	Homo sapiens cDNA clone CS0BJ013YK10	5'-PRIME, mRNA sequence.
VERSION	BX383606_1	GI:30457152	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;		
REFERENCE	1. (bases 1 to 616)		
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
Genoscope - Centre National de Sequenage			
BP 191 91006 EVRY cedex - France			
Email: secref@genoscope.cns.fr			
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9825.r For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DJ013BF05QP1&cluster=9825.r. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DJ013BF05QP1.			
FEATURES	Location/Qualifiers		
source	1. .616		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="BAXON:9605"		
	/clions="CS0DJ013YK10"		
	/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"		
	/cell_line="JURKAT"		
	/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT		

FEATURES	source	Plate: LLM10182 row: i column: 01 High quality sequence stop: 657. 1 . 778
	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:428986" /tissue_type="adenocarcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_91" /note="Organ: Prostate; Vector: PCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	
		Query Match Score 426; DB 12; Length 778; Best Local Similarity 99.7%; Pred. No. 1.7e-204; Indels 1; Gaps 1; Matches 596; Conservative 0; Mismatches 1;
Qy	62	CTTCCTCCCTGCGCTGGCTGTGCTGGTGGAGGTTCCCGAGTCAGATCCCTAAGGAGC 121
Db	26	CTTCCTCCCTGGCTGGCTGGTGGAGGTTCCCGAGTCAGATCCCTAAGGAGC 85
Qy	122	ATGGGCAGCTGATCCATCCCTGGTGTGAACTCTGACTGCAACATGCTGAGCTTA 181
Db	86	ATGGGCAGCTGATCCATCCCTGGTGTGAACTCTGACTGCAACATGCTGAGCTTA 145
Qy	182	CCCAAACCAACACTTAGCTCTCCCTGAAAGATCTCCAGGGCAGAGAGTTCTGGGTCT 241
Db	146	CCCAAACCAACACTTAGCTCTCCCTGAAAGATCTCCAGGGCAGAGAGTTCTGGGTCT 205
Qy	242	CCTAGGACCAAGGACACTGGAGACTTCCAGAAGGGCCCCAACGCCCTAACCTGTCAG 301
Db	206	CCTAGGACCAAGGACACTGGAGACTTCCAGAAGGGCCCCAACGCCCTAACCTGTCAG 265
Qy	302	CCAGGCGATGGTCTAGAGAGCTGCTTCCAGCTTGTGTTGATGACAACCAATTTC 361
Db	266	CCAGGCGATGGTCTAGAGAGCTGCTTCCAGCTTGTGTTGATGACAACCAATTTC 325
Qy	362	TGATGATGATGCTCTGAAGGCTGCTGGAGAACATGGAAACTCTGGCAAGCAGAAA 421
Db	326	TGATGATGATGCTCTGAAGGCTGCTGGAGAACATGGAAACTCTGGCAAGCAGAAA 385
Qy	422	AAATCTCTGCAAGCCAAAGCTTGAATGGTTCTGTCAGGCAAGGCACTGTGACCATG 481
Db	386	AAATCTCTGCAAGCCAAAGCTTGAATGGCTCTGTCAGGCAAGGCACTGTGACCATG 444
Qy	482	GAAGCGAGAGCAAGGCAAGGCCACAGCGTGGCCCTGGAGCTTCCGCCAGGTGGCCG 541
Db	445	GAAGCGAGAGCAAGGCCACAGCGTGGCCCTGGAGCTTCCGCCAGGTGGCCG 504
Qy	542	GCCGAGCTGCGCTGAGACTGGGAGGCATTGACCATGCTCTGAGGATGGAGACTGG 601
Db	505	GCGAGCTGCGCTGAGACTGGGAGGCATTGACCATGCTCTGAGGATGGAGACTGG 564
Qy	602	TGGACGGTGTGCTGTCGAAGGTGTCAGGCAAGAGATTAACATCCCAGGGTCCACAGTGG 659
Db	565	TGGACGGTGTGCTGTCGAAGGTGTCAGGCAAGAGATTAACATCCCAGGGTCCACAGTGG 622
RESULT	8	BG284179
ACCESSION	602408226F1	566 bp tRNA linear EST 21-FEB-2001
DEFINITION	NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520382 5', mRNA sequence.	
KEYWORDS	BG284179	
SOURCE	EST	
ORGANISM	Homo sapiens (human)	

VERSION	BQ053486.1	GI:	19812826	Db	474	TGTTGGACGGTGTCTGAA	494
KEYWORDS	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	RESULT	10				
AUTHORS	1 (bases 1 to 878)	ALB44311	597 bp	mRNA	linear	EST	30-JUL-2002
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	LOCUS	ALB44311 pool_YT_lib_v_SPD	Homo sapiens	cDNA, mRNA sequence.		
JOURNAL	Unpublished (1999)	DEFINITION	ALB44311 pool_YT_lib_v_SPD	Homo sapiens	cDNA, mRNA		
COMMENT	Contact: Robert Straussberg, Ph.D Email: cgsbpsr@mail.nih.gov Tissue Procurement: Dr. Daniel Movicar, DBS/NCI cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: The I.M.A.G.E. Consortium/LINN at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LILCM2122 row: 1 column: 06	ACCESSION	ALB44311				
FEATURES	High quality sequence stop: 394. Location/Qualifiers	VERSION	ALB44311.1	EST			
source	1..878 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:1593553" /tissue_type="natural killer cells, cell line" /lab_host="DH10B (phage-resistant)" /clone_id="NIR_MGC_105" /notes="Organ: blood; Vector: pOTB7; Site_1: XbaI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGCAAG (G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	ORGANISM					
ORIGIN		Query Match	33.0%	Score 390; DB 12; Length 878;			
	Best Local Similarity 99.8%; Pred. No. 3.3e-186; Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Best Local Similarity	100.0%	Score 300; DB 9; Length 597;			
Qy	179 CTACCAAACACATAGCTCCAGGTCCTCGAGATCCTCCAGGTCAGAGTCTGGG	Matches	300;	Score 25.4%; Best Local Similarity 100.0%; Pred. No. 1.4e-140;			
Db	54 CTACCAAACACATAGCTCCAGGTCCTCGAGATCCTCCAGGTCAGAGTCTGGG	Mismatches	0;	Mismatches 0; Indels 0; Gaps 0;			
Qy	239 TGTCCTAGGACAGGACTCCAGACTCCAGCTCCAGGTCAGAGTCTGGG	QY					
Db	54 CTACCAAACACATAGCTCCAGGTCCTCGAGATCCTCCAGGTCAGAGTCTGGG	DB					
Qy	239 TGTCCTAGGACAGGACTCCAGACTCCAGCTCCAGGTCAGAGTCTGGG	DB					
Db	114 TGTCCTAGGACAGGACTCCAGCTCCAGGTCAGAGTCTGGG	QY	884	CGGCTCACCTCCCTCACTCCAGGCTCCATTACTCTGAGCTTGCGGATGAC	943		
Qy	114 TGTCCTAGGACAGGACTCCAGCTCCAGGTCAGAGTCTGGG	DB	37	CGGCTCACCTCCCTCACTCCAGGCTCCATTACTCTGAGCTTGCGGATGAC	96		
Db	114 TGTCCTAGGACAGGACTCCAGCTCCAGGTCAGAGTCTGGG	QY	944	ATCTGCTGCGCTATAAGAGCCGCTCCCTGAGAGGGTGGCCGCTCCCTGAAAG	1003		
Qy	299 CAGCCAGGCAATCGGTCTCAAGAGGTGCTCTCCAAAGCTTGTGACAACCAATT	DB	97	ATCTGCTGCGCTATAAGAGCCGCTCCCTGAGAGGGTGGCCGCTCCCTGAAAG	156		
Db	174 CAGCCAGGCAATCGGTCTCAAGAGGTGCTCTCCAAAGCTTGTGACAACCAATT	QY	1004	GATATACCCCTACTCTGAGCTGCAAGGGAGGAGTCTCTCACTGAGGGTCTCGG	1063		
Qy	359 CCCTCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	DB	157	GATATACCCCTACTCTGAGCTGCAAGGGAGGAGTCTCTCACTGAGGGTCTCGG	216		
Db	234 CCCTCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	QY	1064	TCCCTCCCTGTTTCTGAGCTGCAAGGGAGGAGTCTCTCACTGAGGGTCTCGG	1123		
Qy	419 AGAAAATCTGCCAAGGCCAACCTGAGTTCTCTGAGCCAGGAACTCTGACC	DB	217	TCCCTCCCTGTTTCTGAGCTGCAAGGGAGGAGTCTCTCACTGAGGGTCTCGG	276		
Db	294 AGAAAATCTGCCAAGGCCAACCTGAGTTCTCTGAGCCAGGAACTCTGACC	QY	1124 GAGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1183			
Qy	359 CCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	DB	277 GAGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	336			
Db	354 ATGGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	QY	RESULT	11			
Qy	539 CGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	ALB44307					
Db	414 CGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	LOCUS	ALB44307 pool_AK_lib_v_SPD	Homo sapiens	cDNA, mRNA	614 bp	EST 30-JUL-2002
Qy	599 TGGTGGACGGTGTGTCTGAA	DEFINITION	ALB44307				
Db		ACCESSION	ALB44307				
Qy		VERSION	ALB44307.1	GI:22019089			

KEYWORDS	Homo sapiens (human)	COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgs@bsr@mail.nih.gov Tissue Procurement: James Cleaver, M.D. DNA Library Preparation: Life Technologies, Inc.
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 614)		CDDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN), DNA Sequencing by: Incyte Genomics, Inc.
AUTHORS	Ascroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and Sheridan, P.		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>
TITLE	Homo sapiens EST sequence		Plate: LILN01 Row: d column: 05
JOURNAL	Unpublished (2002)		High Quality sequence stop: 790.
COMMENT	Contact: The Sanger Centre The Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK	FEATURES	Location/Qualifiers
	Email: humquery@sanger.ac.uk	source	1 /db_xref="taxon:9606" /organism="Homo sapiens" /mol_type="mRNA" /clone="IMAGE:4748884" /tissue_type="squamous cell carcinoma" /lab_host="DH10B (T1) Phage-resistant" /clone_lib="NCI CGAP Skn4" /note="Organ: Skin; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI CGAP Library."
	Homo sapiens EST sequence : scd10816.15413A		
	The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool_AK_lib_vSPD cDNA library. Further information can be found at <a href="http://www.sanger.ac.uk/Teams/Team69/">http://www.sanger.ac.uk/Teams/Team69/</a> .	ORIGIN	
	1..614		
	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /map="20" /clone_lib="pool_AK_lib_v_SPD"		
		ORIGIN	
		Query Match Score 276; DB 9; Length 614; Best Local Similarity 100.0%; Pred. No. 2 e-128; Gaps 0; Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
		Qy 908 GCCCTGTCGACCATTACTCTGAGCTGGGATGACATCTGCTGCCCTACTCAAGGAGGCC 967	Qy 929 GAGCTCGGGATGACATCTGCTGCCCTACTCAAGGAGCCCTGCTGAGGAGGCC 988
		Db 44 GCCCTGTCGACCATTACTCTGAGCTGGGATGACATCTGCTGCCCTACTCAAGGAGGCC 103	Db 17 GAGCTCGGGATGACATCTGCTGCCCTACTCAAGGAGCC 76
		Qy 968 TGTGTCCTGAGAGGACTGSCCGTCCCTGGCAAGGATAAACCTCTACCTGCACTTG 1027	Qy 989 CGCTCTCTGGAAAGATAACCCCTACCTGCACTTG 1048
		Db 104 TGTGTCCTGAGAGGCTGGCCGTCCCTGGCAAGGATAACCCCTACCTGCACTTG 163	Db 77 CGCTCTCTGGAAAGATAACCCCTACCTGCACTTG 136
		Qy 1028 CAGAGGACACCACTCAACTGAAAGAGCTGACAGCTCTCCCTCCCTGTTCTGAAGGCTGCC 1087	Qy 1049 AAAGAGCTGGAGCTGGCAGCTCCCTGCTGAACTGCTGCAAGGGAGGAGGCTCTCTC 1108
		Db 164 CAGAGGACACCACTCAACTGAAAGAGCTGACAGCTCTCCCTCCCTGTTCTGAAGGCTGCC 223	Db 137 AAGAGCTGGAGCTGGCAGCTCCCTGCTGAACTGCTGCAAGGGAGGAGGCTCTCTC 196
		Qy 1088 ACAGGGGAGGAGTCTCTCACTGAGGGCTCCCTGGAGTCTCATGAGC 1147	Qy 1109 AGTGGGGCTCTGGGGAGTCCCTCAAGCTTCTAACATGAGGCTCTCTCT 1168
		Db 224 ACAGGGGAGGAGTCTCTCACTGAGGGCTCCCTGGAGTCTCATGAGC 283	Db 197 AGTGGGGCTCTGGGGAGTCCCTCAAGCTTCTAACATGAGGCTCTCTCT 256
		Qy 1148 CTGAATGACAGGGCTCTCTCACTGAGGGCTCCCTGGAGTCTCATGAGC 319	Qy 1169 TGGATGATGCTAG 1183
		Db 284 CTGAATGACAGGGCTCTCTCACTGAGGGCTCCCTGGAGTCTCATGAGC 283	Db 257 TTGGATGATGCTAG 271
RESULT	12	RESULT 13	
LOCUS	BG677567	LOCUS BU944126	
DEFINITION	BG677567_602624118F1 NCI CGAP_Skn4 mRNA Homo sapiens cDNA clone IMAGE:4748884_5', mRNA sequence.	DEFINITION AGECOUNT 10545003 NIH_MGC_107_Homo_sapiens_CDNA_clone IMAGE:6728350_5', mRNA sequence.	
ACCESSION	BG677567	ACCESSION BU944126	
VERSION	1	VERSION BU944126_1 GI:24132945	
KEYWORDS		KEYWORDS EST	
SOURCE	Homo sapiens (human)	SOURCE Homo sapiens (human)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 7567)	REFERENCE 1 (bases 1 to 960) NIH-MGC http://mgc.ncbi.nih.gov/ Unpublished (1999)	
AUTHORS		AUTHORS Contact: Robert Strausberg, Ph.D. Email: cgs@bsr@mail.nih.gov	
TITLE		TITLE Tissue Procurement: AACC CDNA Library Preparation: Rubin Laboratory DNA Sequencing by: Agencourt Bioscience Corporation	
JOURNAL		JOURNAL Unpublished (1999)	

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINN at:  
<http://image.llnl.gov>  
 Plate: LUCM349 row: M column: 21  
 High quality sequence stop: 628.  
 Location/Qualifiers  
 1. 960  
   /organism="Homo sapiens"  
   /mo\_type="mRNA"  
   /db\_xref="taxon:9606"  
   /clone="IMAGE:6728350"  
   /tissue\_type="adenocarcinoma, cell line"  
   /lab\_host="DHLB (phage-resistant)"  
   /clone\_id="NHG MGC 107"  
   /noe="Organ: breast; Vector: pORN7; Site\_1: EcoRI;  
 Site\_2: XbaI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XbaI sites using the  
 following 5' adaptor: GCAACGA(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

FEATURES	at <a href="http://www.sanger.ac.uk/Teams/Team69/">http://www.sanger.ac.uk/Teams/Team69/</a> .
source	/location/Qualifiers 1. . 606
	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /map="20" /clone_lib="pool_FLU_lib_v_SPC"
ORIGIN	
	Query Match 16.9%; Score 200; DB 9; Length 60;
	Best Local Similarity 99.6%; Pred. No. 8.2e-30; Indels 0; Gaps 0;
	Matches 250; Conservative 0; Mismatches 1;
QY	908 GCCTGGGACATTACTCGAGCAGGGATGACATCTGGCTACTAAGGCC 967
Db	46 Gccctgggaccattactcgagcaggatgtacatcggtggatgacatctggctactaaggcc 105
QY	968 TGTGTCCTGCAGAGGGCTGGCCGCTOCTGCGCAAGGATAACCCCTAACGTTGACTGTG 1027
Db	106 Tgtgtccctgcagaggctggccgtctggactgtgttttcatatcaggcc 165
ORIGIN	
	Query Match 16.9%; Score 200; DB 9; Length 60;
	Best Local Similarity 99.6%; Pred. No. 8.2e-30; Indels 0; Gaps 0;
	Matches 250; Conservative 0; Mismatches 1;
QY	908 GCCTGGGACATTACTCGAGCAGGGATGACATCTGGCTACTAAGGCC 967
Db	46 Gccctgggaccattactcgagcaggatgtacatcggtggatgacatctggctactaaggcc 105
QY	968 TGTGTCCTGCAGAGGGCTGGCCGCTOCTGCGCAAGGATAACCCCTAACGTTGACTGTG 1027
Db	106 Tgtgtccctgcagaggctggccgtctggactgtgttttcatatcaggcc 165
ORIGIN	
	Query Match 17.3%; Score 205; DB 13; Length 960;
	Best Local Similarity 100.0%; Pred. No. 2.8e-92; Indels 0; Gaps 0;
	Matches 205; Conservative 0; Mismatches 0;
QY	979 GAGGAGTGCGCTGCTGGCAAGGATTAACCCCTACCMGTGACTGTGCAAGGACACC 1038
Db	177 GAGGGCTGGCGCTCCCTGGAAAGGATTAACCCCTACCMGTGACTGTGCAAGGACACC 236
QY	1039 ACTCAACTGGAGAGGTGGAGCTCCCTCTGTTCTGAAAGTGCACAGGGAGA 1098
Db	237 ACTCAACTGGAGAGGTGGAGCTCCCTCTGTTCTGAAAGTGCACAGGGAGA 2956
QY	1099 GTCGTCTCTAGTAGGGATGATGCCCTAGCTCTACATCGCTGAATGACCA 1158
Db	297 GTCGTCTCTAGTAGGGATGCCCTAGCTCTACATCGCTGAATGACCA 356
QY	1159 GGTTGTCCTCTGGATGATGCCCTAG 1183
Db	357 GGTTGTCCTCTGGATGATGCCCTAG 381
RESULT 14	
AL844309	606 bp mRNA linear EST 30-JUL-2002
LOCUS	AL844309 pool_FLU_lib_v_SPC Homo sapiens cDNA, mRNA sequence.
DEFINITION	
VERSION	AL844309.1 GI:22019091
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE	1 (bases 1 to 606)
AUTHORS	Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and Sheridan, E.
TITLE	
JOURNAL	The Sanger Centre
COMMENT	The Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK Email: hungry@sanger.ac.uk Sanger Centre name : sccd10816.400489A Sanger Centre name : sccd10816.400489A Homo sapiens EST sequence Homo sapiens EST sequence Unpublished (2002) Contact: The Sanger Centre The Sanger Centre The Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK Email: hungry@sanger.ac.uk Sanger Centre name : sccd10817.154136A Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool_AK_lib_v_SPC cDNA library. Further information can be found at <a href="http://www.sanger.ac.uk/Teams/Team69/">http://www.sanger.ac.uk/Teams/Team69/</a> .
FEATURES	
source	1. . 612
	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /map="20" /clone_lib="pool_AK_lib_v_SPC"
ORIGIN	

Query Match 9.8%; Score 116; DB 9; Length 642;  
Best Local Similarity 100.0%; Pred. No. 3.7e-47;  
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 116;保守性 0; 差错 0; 插入 0; 缺失 0;

Qy	Db	Score	Length
1068	588	TCCCTGTTTCCTGAAAGTGCACAGGGAGTCCTGAGGTCTCTCACTGAGGCTCTCGGGAGT	1127
Db	528	CCCTCAGCTCTACATCAGCCTGAATGACAGGGCTGCTCTTGATGATGCCAGT	1183
Qy	528	CCCTCAGCTCTACATCAGCCTGAATGACAGGGCTGCTCTTGATGATGCCAGT	473

Search completed: February 20, 2004, 08:13:51  
Job time : 2937 secs

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 05:30:56 i Search time 113 Seconds  
(without alignments) Seconds

5809.796 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agttagactccaaaggacc.....ttctttggatgtccatg 1183

Scoring table: OLIGO\_NTUC

Gapop 60.0 , Gapext 60.0

Searched: 682769 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
 1: /cgnd\_6\_ptodata/2/ina/5A-COMB.seq: \*  
 2: /cgnd\_6\_ptodata/2/ina/5B-COMB.seq: \*  
 3: /cgnd\_6\_ptodata/2/ina/6A-COMB.seq: \*  
 4: /cgnd\_6\_ptodata/2/ina/6B-COMB.seq: \*  
 5: /cgnd\_6\_ptodata/2/ina/PC7US\_COMB.seq: \*  
 6: /cgnd\_6\_ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	20	1.7	539	4	US-09-621-976-10381	
2	20	1.7	675	1	US-08-707-733A-3	
3	20	1.7	675	1	US-08-707-732A-3	
4	20	1.7	2129	4	US-09-145-1452	
5	20	1.7	786431	4	US-09-751-339-3	
6	19	1.6	1467	4	US-09-579-132-2	
7	19	1.6	1548	4	US-09-099-053-1	
8	19	1.6	2771	4	US-19-016-434-1101	
9	18	1.5	1438	3	US-09-187-314-4	
10	18	1.5	1438	4	US-09-170-966-4	
11	18	1.5	1669	3	US-09-318-448-8	
c	12	18	1.5	3050	3	US-09-216-511-78
c	13	18	1.5	700000	4	US-09-851-890-3
c	14	17	1.4	351	3	US-09-016-432-1
c	15	17	1.4	351	4	US-08-822-89-C-1
c	16	17	1.4	351	4	US-09-008-80A-3
c	17	17	1.4	351	4	US-09-404-417A-1
c	18	17	1.4	435	4	US-09-52-97A-6817
c	19	17	1.4	439	4	US-09-222-57-172
c	20	17	1.4	439	4	US-09-889-61-172
c	21	17	1.4	439	4	US-09-620-45B-172
c	22	17	1.4	439	4	US-09-339-3-172
c	23	17	1.4	439	4	US-09-333-86B-172
c	24	17	1.4	439	4	US-09-604-28A-172
c	25	17	1.4	439	4	US-09-25-480-172
c	26	17	1.4	439	4	US-09-834-75-172
c	27	17	1.4	445	4	US-09-702-705-1598

28 17 1.4 445 4 US-09-736-457-1598  
 29 17 1.4 445 4 US-09-614-124B-1598  
 30 17 1.4 445 4 US-09-671-328-1598  
 31 17 1.4 541 4 US-09-220-132-10  
 32 17 1.4 566 4 US-09-671-976-1574  
 c 33 17 1.4 627 4 US-09-328-352-1086  
 34 17 1.4 643 4 US-09-833-381-1262  
 c 35 17 1.4 651 4 US-09-016-434-1255  
 36 17 1.4 674 4 US-09-621-976-87  
 37 17 1.4 759 4 US-09-493-039A-5493  
 38 17 1.4 894 4 US-09-334-354-1  
 c 39 17 1.4 941 4 US-09-634-230-89  
 c 40 17 1.4 1089 1 US-09-154-915-1  
 41 17 1.4 1089 2 US-08-464-51-37  
 42 17 1.4 1089 2 US-08-453-367A-37  
 43 17 1.4 1089 5 PCU-US9-09945-1  
 44 17 1.4 1089 5 Sequence 37, Appli  
 45 17 1.4 1747 4 Sequence 37, Appli  
 Sequence 66, Appli

ALIGNMENTS

RESULT 1  
 US-09-621-976-10381  
*i* Sequence 10381 Application US/09621976  
*i* Patent No. 6639063  
*i* GENERAL INFORMATION:  
*i* APPLICANT: Dumas Milne Edwards, J.B.  
*i* ATTORNEY: Jobert, S.  
*i* TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
*i* FILE REFERENCE: GENSET 054P22  
*i* CURRENT APPLICATION NUMBER: US/09/621.976  
*i* CURRENT FILING DATE: 2000-07-21  
*i* NUMBER OF SEQ ID NOS: 19335  
*i* SOFTWARE: Patent.pml  
*i* SEQ ID NO: 10381  
*i* LENGTH: 539  
*i* TYPE: DNA  
*i* ORGANISM: Homo sapiens  
 US-09-621-976-10381

Query Match 1.7%; Score 20; DB 4; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 472 TGTGACCATGGAGCAGAGGA 491  
 Db 52 TGTGACCATGGAGCAGAGGA 71

RESULT 2  
 US-08-707-799A-3  
*i* Sequence 3, Application US/08707793A  
*i* Patent No. 5776696  
*i* GENERAL INFORMATION:  
*i* APPLICANT: SALOME, SCOTT P.  
*i* TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING  
*i* CORRESPONDENCE ADDRESS:  
*i* ADDRESS: Meck & Co., Inc.  
*i* STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
*i* CITY: Railway  
*i* STATE: NJ  
*i* COUNTRY: USA  
*i* ZIP: 07065-0900  
*i* COMPUTER READABLE FORM:  
*i* MEDIUM TYPE: Diskette  
*i* COMPUTER: IBM Compatible  
*i* OPERATING SYSTEM: DOS  
*i* SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/707,793A  
 FILING DATE: 04-SEP-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Camara, Valerie J  
 REGISTRATION NUMBER: 35,090  
 REFERENCE/DOCKET NUMBER: 19494  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 908-594-3902  
 TELEX: 908-594-4720  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 675 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: Genomic DNA

US-08-707-793A-3

Query Match 3      Score 20; DB 1; Length 675;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 20; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy      747 CCTTCCTCATCGGGAGAC 766  
 Db      443 CCTTCCTCATCGGGAGAC 462

RESULT 3  
 US-08-707-792A-3  
 Sequence 3, Application US/08/707,792A  
 Patent No. 5733398  
 GENERAL INFORMATION:  
 APPLICANT: MARCY, ALICE  
 APPLICANT: SALOME, SCOTT P.  
 APPLICANT: WISNIEWSKI, DOUGLAS  
 TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING  
 TITLE OF INVENTION: FUSION PROTEINS  
 NUMBER OF SEQUENCES: 17  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Merck & Co., Inc.  
 STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
 CITY: Rahway  
 STATE: NJ  
 COUNTRY: USA  
 ZIP: 07065-0900  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/707,792A  
 FILING DATE: 04-SEP-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Camara, Valerie J  
 REGISTRATION NUMBER: 35,090  
 REFERENCE/DOCKET NUMBER: 19524  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 908-594-3902  
 TELEX: 908-594-4720  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:

; LENGTH: 675 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; US-08-707-792A-3

Query Match 4      Score 20; DB 1; Length 675;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 20; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy      747 CCTTCCTCATCGGGAGAC 766  
 Db      443 CCTTCCTCATCGGGAGAC 462

RESULT 4  
 US-09-016-434-1452  
 Sequence 4, Application US/09/016434  
 Patent No. 6500938  
 GENERAL INFORMATION:  
 APPLICANT: Jeffrey J. Seilhamer  
 APPLICANT: Jeffrey J. Seilhamer  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 NUMBER OF SEQUENCES: 1490  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 For Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/016,434  
 FILING DATE: HEREWITH  
 CLASSIFICATION NUMBER: US/09/016,434  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0002 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 845-4166  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1452:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2129 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: 9775207  
 US-09-016-434-1452

Query Match 5      Score 20; DB 4; Length 2129;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 20; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy      747 CCTTCCTCATCGGGAGAC 766  
 Db      508 CCTTCCTCATCGGGAGAC 527

RESULT 5  
US-09-751-389-3  
Sequence 3, Application US/09751389

; Patent No. 650628

; GENERAL INFORMATION:

; APPLICANT: GUEGLIER, Karl et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001067

; CURRENT APPLICATION NUMBER: US/09/751,389

; CURRENT FILING DATE: 2001-01-02

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 786431

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: n = A,T,C or G

US-09-751-389-3

; Query Match

; Best Local Similarity 1.7%; Score 20; DB 4; Length 786431;

; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 693 TGACCAAGGAAACAGAG 712

Db 412751 TGACCAAGGAAACAGAG 412770

RESULT 6  
US-09-579-182-2

; Sequence 2, Application US/09579182

; Patent No. 650628

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: NUCLEAR ACID MOLECULES ENCODING HUMAN KINASE AND

; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR

; FILE REFERENCE: MNI-161

; CURRENT APPLICATION NUMBER: US/09/579,182

; CURRENT FILING DATE: 2000-05-25

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1467

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-579-182-2

; Query Match

; Best Local Similarity 1.6%; Score 19; DB 4; Length 1467;

; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 AGGGGCCTTCATCCGG 760

Db 423 AGGGGCCTTCATCCGG 441

RESULT 7  
US-09-099-053-1

; Sequence 1, Application US/09099053

; Patent No. 6588063

; GENERAL INFORMATION:

; APPLICANT: Greg Plowman

; APPLICANT: Susan Orntust

; APPLICANT: David Markby

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF

; SAD RELATED DISORDERS

; NUMBER OF SEQUENCES: 28

Qy 742 AGGGGCCTTCATCCGG 760

Db 423 AGGGGCCTTCATCCGG 441

RESULT 8  
US-09-099-053-1

; Sequence 1, Application US/09099053

; Patent No. 6500938

; GENERAL INFORMATION:

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

; TITLE OF INVENTION: PATHWAY GENE EXPRESSION

; NUMBER OF SEQUENCES: 1490

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/016,434

; FILING DATE: HERMIT

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

;

;

;

;

;

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Pritch Street  
CITY: Los Angeles  
STATE: California  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPILER: Storage  
OPERATING SYSTEM: IBM PC Compatible  
SOFTWARE: FastSEQ for Windows 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/099,053  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/049,914  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 235/121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1548 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-099-053-1

Query Match 1.6%; Score 19; DB 4; Length 1548;  
Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 742 AGGGGCCTTCATCCGG 760  
Db 471 AGGGGCCTTCATCCGG 489

RESULT 8  
US-09-099-053-1

Patent No. 6500938

GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HERMIT  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
;

APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37, 071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2771 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: 91256002  
US-09-016-434-1101

Query Match Score 1.6%; Pred. No. 12; Length 2771;  
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

RESULT 9  
US-09-187-331-4  
Sequence 4, Application US/09187331  
Patent No. 6043056

GENERAL INFORMATION:  
APPLICANT: Yue, Henry  
APPLICANT: Corley, Neil C.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Gorgone, Gina A.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: FILE REFERENCE: FF-0631 US  
CURRENT APPLICATION NUMBER: US/09/187,331  
CURRENT FILING DATE: 1998-11-06  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PERL Program  
SEQ ID NO 4  
LENGTH: 1438  
TYPE: DNA  
FEATURE: -  
OTHER INFORMATION: 2705267  
US-09-187-331-4

Query Match Score 1.5%; Pred. No. 38; Length 1438;  
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

RESULT 10  
US-09-470-946-4  
Sequence 4, Application US/09470946  
Patent No. 6358923  
GENERAL INFORMATION:  
APPLICANT: Yue, Henry  
APPLICANT: Corley, Neil C.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Gorgone, Gina A.  
APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS  
FILE REFERENCE: FF-0631 US  
CURRENT APPLICATION NUMBER: US/09/470,946  
CURRENT FILING DATE: 1999-12-22  
EARLIER APPLICATION NUMBER: US 09/187,331  
EARLIER FILING DATE: 1998-11-06  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PERL Program  
SEQ ID NO 4  
LENGTH: 1438  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE: -  
OTHER INFORMATION: 2705267  
US-09-470-946-4

Query Match Score 1.5%; Pred. No. 38; Length 1438;  
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

RESULT 11  
US-09-318-448-8  
Sequence 8, Application US/09318448  
Patent No. 6210950

GENERAL INFORMATION:  
APPLICANT: Johnson, William G.  
APPLICANT: Stenicos, Edward S.  
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
TITLE OF INVENTION: DEVELOPMENTAL DISORDERS  
FILE REFERENCE: 601-1-057

CURRENT APPLICATION NUMBER: US/09/318,448  
CURRENT FILING DATE: 1999-05-25  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 1669  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-318-448-8

Query Match Score 1.5%; Pred. No. 38; Length 1669;  
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

RESULT 12  
US-09-276-531-78/C  
Sequence 78, Application US/09276531  
Patent No. 6183968

GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
APPLICANT: Reddy, Roopa  
APPLICANT: Guegler, Karl J.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING  
TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION  
NUMBER OF SEQUENCES: 134  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO

STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/276,531  
 FILING DATE: Herewith  
 CLASSIFICATION:  
 APPLICATION NUMBER: 60/079,677  
 FILING DATE: March 27, 1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lynn E. Murry, Ph.D.  
 REGISTRATION NUMBER: 42,918  
 REFERENCE/DOCKET NUMBER: PA-0008 US  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 78:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 309 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: BRAINOT14  
 CLONE: 1535762  
 US-09-276-531-78

Query Match Score: 1.5% ; Length: 3090;  
 Best Local Similarity: 100.0% ; Pred. No. 38; Mismatches: 0; Indels: 0; Gaps: 0;

Qy 1093 GGAGGAGTCCTCTTCAG 1110  
 Db 398 GGAGGAGTCCTCTTCAG 381

RESULT 13  
 US-09-851-896-3  
 Sequence 3, Application US/09851896  
 Patent No. 6410325  
 GENERAL INFORMATION:  
 APPLICANT: C. Frank Bennett  
 APPLICANT: Susan M. Freier  
 APPLICANT: Andrew T. Watt  
 TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT EXPRESSION)  
 FILE REFERENCE: RTS-0220  
 CURRENT APPLICATION NUMBER: US/09/851,896  
 CURRENT FILING DATE: 2001-05-08  
 NUMBER OF SEQ ID NOS: 89  
 SEQ ID NO 3  
 LENGTH: 70000  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 US-09-851-896-3

Query Match Score: 1.5% ; Length: 70000;  
 Best Local Similarity: 100.0% ; Pred. No. 36; Mismatches: 0; Indels: 0; Gaps: 0;

Qy 455 GTCCAAAGGCCAGGGACT 472  
 Db 60708 GTCCAAAGGCCAGGGACT 60725

RESULT 14  
 US-09-046-479-1/C  
 Sequence 1, Application US/09046479  
 Patent No. 6291633  
 GENERAL INFORMATION:  
 APPLICANT: Sheppard, Paul O.  
 APPLICANT: Deisher, Theresa A.  
 TITLE OF INVENTION: MOTILIN HOMOLOGS  
 NUMBER OF SEQUENCES: 7  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ZymoGenetics Inc.  
 STREET: 1201 Eastlake Avenue East  
 CITY: Seattle  
 STATE: WA  
 COUNTRY: USA  
 ZIP: 98102  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSQL for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/046,479  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sawislak, Deborah A.  
 REGISTRATION NUMBER: 37,138  
 REFERENCE/DOCKET NUMBER: 97-04  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-442-6672  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 351 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: Coding Sequence  
 LOCATION: 1..351  
 OTHER INFORMATION:  
 NAME/KEY: sig\_peptide  
 LOCATION: 1..69  
 OTHER INFORMATION:  
 NAME/KEY: mat\_peptide  
 LOCATION: 70..351  
 OTHER INFORMATION:  
 US-09-046-479-1

Query Match Score: 1.4% ; Length: 351;  
 Best Local Similarity: 100.0% ; Pred. No. 1.2e+02; Mismatches: 0; Indels: 0; Gaps: 0;

Qy 296 GTCCAAAGGCCAGGGACTC 312  
 Db 57 GTCCAAAGGCCAGGGACTC 41

RESULT 15  
 US-08-822-897C-1/C  
 Sequence 1, Application US/08822897C  
 Patent No. 6380158  
 GENERAL INFORMATION:  
 APPLICANT: Sheppard, Paul O.  
 APPLICANT: Deisher, Theresa A.  
 TITLE OF INVENTION: MOTILIN HOMOLOGS  
 NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ZymoGenetics, Inc.  
 STREET: 1201 Eastlake Avenue East  
 CITY: Seattle  
 STATE: WA  
 COUNTRY: USA  
 ZIP: 98102

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/822,897C  
 FILING DATE:  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Swislik, Deborah A.  
 REGISTRATION NUMBER: 37,438  
 REFERENCE/DOCKET NUMBER: 97-04  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-442-6672  
 TELEFAX: 206-442-6678  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 351 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

FEATURE:  
 NAME/KEY: Coding Sequence  
 LOCATION: 1...351  
 OTHER INFORMATION:  
 NAME/KEY: sig\_peptide  
 LOCATION: 1...69  
 OTHER INFORMATION:  
 NAME/KEY: mat\_peptide  
 LOCATION: 70...351  
 OTHER INFORMATION:  
 US-08-822-897C-1

Query Match 1.4%; Score 17; DB 4; Length 351;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 296 GTCCAGCCAGGATGC 312  
 Db 57 GTCCAGCCAGGATGC 41

Search completed: February 20, 2004, 08:16:02  
 Job time : 121 secs

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Inc.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 07:26:02 ; Search time 486 Seconds  
(without alignments)

8623.550 Million cell updates/sec

Title: US-09-939-853A-74

Percent score: 1183

Sequence: 1 agcttagtccaaaggacc.....ttctttggatgtgcctag 1183

Scoring table: OLIGO\_NUC

Gapext 60.0 , Gapext 60.0

Searched: 2308684 seqs, 1750822206 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0

Maximum DB seq length: 0

Post-processing: Listing first 45 summaries

Database : 1: /cgns\_6/\_ptodata/2/pubpna/us07\_PUBCOMB.seq; \*  
2: /cgns\_6/\_ptodata/2/pubpna/dctm\_new\_pub.seq; \*  
3: /cgns\_6/\_ptodata/2/pubpna/us06\_new\_pub.seq; \*  
4: /cgns\_6/\_ptodata/2/pubpna/us05\_new\_pub.seq; \*  
5: /cgns\_6/\_ptodata/2/pubpna/us07\_new\_pub.seq; \*  
6: /cgns\_6/\_ptodata/2/pubpna/pcmts\_pubcomb.seq; \*  
7: /cgns\_6/\_ptodata/2/pubpna/us08\_new\_pub.seq; \*  
8: /cgns\_6/\_ptodata/2/pubpna/us09\_new\_pub.seq; \*  
9: /cgns\_6/\_ptodata/2/pubpna/us09a\_pubcomb.seq; \*  
10: /cgns\_6/\_ptodata/2/pubpna/us09b\_pubcomb.seq; \*  
11: /cgns\_6/\_ptodata/2/pubpna/us09c\_pubcomb.seq; \*  
12: /cgns\_6/\_ptodata/2/pubpna/us09\_new\_pub.seq; \*  
13: /cgns\_6/\_ptodata/2/pubpna/us10\_pubcomb.seq; \*  
14: /cgns\_6/\_ptodata/2/pubpna/us10c\_pubcomb.seq; \*  
15: /cgns\_6/\_ptodata/2/pubpna/us10c\_pubcomb.seq; \*  
16: /cgns\_6/\_ptodata/2/pubpna/us10\_new\_pub.seq; \*  
17: /cgns\_6/\_ptodata/2/pubpna/us10c\_new\_pub.seq; \*  
18: /cgns\_6/\_ptodata/2/pubpna/us60\_pubcomb.seq; \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	735	62.1	786	14	US-10-043-649-1
2	724	61.2	864	10	US-09-84-353-21302
3	657	55.5	763	9	US-09-867-550-953
4	348	29.4	444	9	US-09-867-550-951
5	341	28.8	875	9	US-09-867-550-1915
6	134	11.3	432	9	US-09-864-761-2829
7	134	11.3	448	9	US-09-864-761-15513
8	96	8.1	96	9	US-09-864-761-19612
9	87	7.4	320	10	US-09-84-353-17314
10	77	6.5	152	10	US-09-814-353-4631
11	77	6.5	152	10	US-09-814-353-10930
c 12	21	1.8	701	15	US-10-062-674-2188
c 13	20	1.7	611	15	US-10-057-632-19582
14	20	1.7	672	15	US-10-027-632-107077
15	20	1.7	672	15	US-10-027-632-142058

Sequence 142059,  
Sequence 142060,  
Sequence 26286, A  
Sequence 640, APP  
Sequence 1776, APP  
Sequence 27, Appl  
Sequence 934, APP  
Sequence 1452, APP  
Sequence 3, Appl  
Sequence 4510, APP  
Sequence 30106, A  
Sequence 237, APP  
Sequence 13665, A  
Sequence 20, Appl  
Sequence 55347, A  
Sequence 11, Appl  
Sequence 25, Appl  
Sequence 1, Appl  
Sequence 577, APP  
Sequence 1101, APP  
Sequence 46381, A  
Sequence 2448, APP  
Sequence 321, APP  
Sequence 8009, APP  
Sequence 1885, APP  
Sequence 195991,  
Sequence 26739, APP  
Sequence 38019, APP  
Sequence 270409, APP

## ALIGNMENTS

RESULT 1	US-10-043-649-1	Sequence 1, Application US/10043649 ; Publication No. US20030053924A1 ; GENERAL INFORMATION:
		; APPLICANT: Holland, Sacha J.
		; Mendenhall, Marcy K.
		; Pardo, Jorge
		; Spencer, Collin
		; FU, C. Alan
		; LUO, Ying
		; PAYAN, Donald G.
		; MANCEBO, Ning
		; WU, Jun
		; ZHOU, Xiulan
		; SHEN, Mary
		; LIAO, X. Charlene
		; APPLICANT: Sheng, Ning
		; TITLE OF INVENTION: Cloning of a No. US200305994A1 Inhibitor of Antigen-receptor
		; FILE REFERENCE: A-70219-1/RMS/DHR
		; CURRENT APPLICATION NUMBER: US/10/043, 649
		; CURRENT FILING DATE: 2002-01-10
		; PRIOR APPLICATION NUMBER: US 60/260, 953
		; PRIOR FILING DATE: 2001-01-10
		; NUMBER OF SEQ ID NOs: 3
		; SOFTWARE: PatentIn version 3.1
		; SEQ ID NO: 1
		; LENGTH: 786
		; TYPE: DNA
		; ORGANISM: Homo sapiens
		; FEATURE: NAME/KEY: CDS
		; LOCATION: (1). (786)
		; OTHER INFORMATION: US-10-043-649-1
		Query Match
		Score 735; DB 14; Length 786;

CURRENT FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: US 60/191,031  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: US 60/207,124  
 PRIOR FILING DATE: 2000-05-25  
 PRIOR APPLICATION NUMBER: US 60/211,940  
 PRIOR FILING DATE: 2000-06-15  
 PRIOR APPLICATION NUMBER: US 60/216,820  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: US 60/220,661  
 PRIOR FILING DATE: 2000-07-25  
 PRIOR APPLICATION NUMBER: US 60/257,672  
 PRIOR FILING DATE: 2000-12-21  
 NUMBER OF SEQ ID NOS: 22037  
 SOFTWARE: FASSEQ for Windows Version 4.0  
 SEQ ID NO: 213302  
 LENGTH: 864  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 1..2, 3..32, 862..863, 864  
 OTHER INFORMATION: n = A,T,C or G  
 3-09-814-353-213.02

Query Match 61.2%; Score 724;  
 Best Local Similarity 99.9%; Pred. No. 0  
 Matches 774; Conservative 0; Mismatch 2

Y	2	GCTAGAGCTCCAAAGGACCCCCAACGGCTGTGTC
Y	54	GETTAGAGCTCCAAAGGCCCAAGCTGTGTC
Y	62	CCTCCCTCCCTGGCTGGCTTGCTTGGAG
Y	114	CCTCCCTCCCTGGCTGGCTGGCTTGGAG
Y	122	AATGGGAGGTGATCCATCCTGGTACA
Y	174	AATGGGAGGTGATCCATCCTGGTACA
Y	182	CCCAAAACRAAACCTAGCCCTTCCTGAAGA
Y	234	CCCAAAACRAAACCTAGCCCTTCCTGAAGA
Y	242	CTTAGGACCAAGGACACTGGAGACTTCCAG
Y	294	CTTAGGACCAAGGACACTGGAGACTTCCAG
Y	302	CCAGAGATGGCTCTGAGAGCTGTCTGAG
Y	354	CCAGAGATGGCTCTGAGAGCTGTCTGAG
Y	362	TGATGATGATGTTCTGAGCTGTCTGAG
Y	414	TGATGATGTTCTGAGCTGTCTGAG
Y	422	AAATCTCTGCAGGCCAACCTGGTCTCTCT
Y	474	AAATCTCTGCAGGCCAACCTGGTCTCTCT
Y	482	GAAGCAGAGAAGGAGCCACACCGTGTG
Y	534	GAAGCAGAGAAGGAGCCACACCGTGTG
Y	542	GGCGAGCTGTCTGAGACTGGGAGAGCAT
Y	594	GGCGAGCTGTCTGAGACTGGGAGAGCAT
Y	602	TGACAGCTGTCTGAGACTGGGAGAGCAT
Y	654	TGACAGCTGTCTGAGACTGGGAGAGCAT

**RESULT 3**

US-09-867-550-953  
 Sequence 953, Application US/09867550  
 Patent No. US20020082206A1

GENERAL INFORMATION:  
 APPLICANT: Leach, Martin D.  
 APPLICANT: Mehraban, Fuaad,  
 APPLICANT: Conley, Pamela  
 APPLICANT: Law, Debbie  
 APPLICANT: Topper, James  
 TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and  
 FILE REFERENCE: 21402-013 (Curia-313)

CURRENT FILING DATE: 2001-09-20  
 PRIOR APPLICATION NUMBER: US/09/867,550  
 PRIORITY DATE: 2000-05-30  
 NUMBER OF SEQ ID NOS: 2125  
 SOFTWARE: Fast-SEQ for Windows Version 4.0  
 SEQ ID NO: 953  
 LENGTH: 763  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-867-550-953

Query Match Score 55.5%; Score 657; DB 9; Length 763;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 117 GAGGATGGGAGGTATCCATCCGTGTACAAACTCTGTGACTGCAGACAGATGCTG 176  
 5 GAGGCAAGGGGAGGTATCCATCCGTGTACAAACTCTGTGACTGCAGACAGATGCTG 64  
 Qy 177 AGCTTACCAACCACCTTACCTTCCTGAAGATCTCCCTGAAGCTCCCTAACCTG 236  
 Db 65 AGCTTACCAACCACCTTACCTTCCTGAAGATCTCCCTAACCTG 124  
 Qy 237 GGTTGTCCTAGACCAAGGAACTGGCAAGTTCAGAAGCTCCCTAACCTG 236  
 Db 125 GATGTCCTAGACCAAGGAACTGGCAAGTTCAGAAGCTCCCTAACCTG 184  
 Qy 297 TCCAGCGACAGCATGGTCAAGCAGCTCTTCCCAAGCTTGTGACAAACCAT 356  
 Db 185 TCAAGCGACAGCATGGTCAAGCAGCTCTTCCCAAGCTTGTGACAAACCAT 244  
 Qy 357 TTCCCTCGATGATGATGCTCTTCTGAGCTCTCTGCTGAGGAACTCTGCCAGCA 416  
 Db 245 TTCCCTCGATGATGATGCTCTTCTGAGCTCTCTGCTGAGGAACTCTGCCAGCA 304  
 Qy 417 GAGGAAATCTCTGCCAAAGCTTCAAGCTTCAAGCTTCAAGCTTGTGA 476  
 Db 305 CAAGAAATCTCTGCCAAAGCTTCAAGCTTCAAGCTTCAAGCTTGTGA 364  
 Qy 477 CCATGAAAGCAGAGGAAAGGAAAGGAAAGGAAAGCTGGGGAGCTGGCTCTGAGGATGGAG 536  
 Db 365 CCATGAAAGCAGAGGAAAGGAAAGGAAAGGAAAGCTGGGGAGCTGGCTCTGAGGATGGAG 424  
 Qy 537 CCCTGACCCAGCTCTGAGCTGGGGAGCCATTGACCATGCTCTGAGGATGGAG 596  
 Db 425 GCGGAGCTGCTGGCTGAGCTGGGGAGCTGGCTCTGAGGATGGAG 484  
 Qy 597 ACTGGTGGACGGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 656  
 Db 485 ACTGGTGGACGGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 544

RESULT 4

US-09-867-550-951  
 Sequence 951, Application US/09867550  
 Patent No. US20020082206A1

GENERAL INFORMATION:  
 APPLICANT: Leach, Martin D.  
 APPLICANT: Mehraban, Fuaad,  
 APPLICANT: Conley, Pamela  
 APPLICANT: Law, Debbie  
 APPLICANT: Topper, James  
 TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and  
 FILE REFERENCE: 21402-013 (Curia-313)

CURRENT FILING DATE: 2001-09-20  
 PRIOR APPLICATION NUMBER: US/09/867,550  
 NUMBER OF SEQ ID NOS: 2125  
 SOFTWARE: Fast-SEQ for Windows Version 4.0  
 SEQ ID NO: 951  
 LENGTH: 444  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-867-550-951

Query Match Score 29.4%; Score 348; DB 9; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-167;  
 Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 CCTAGGACCAAGGAACTGGCAAGCTTCAGAAGGGCCCCAAAGCCTAACCTGTCCAG 301  
 Qy 302 CCAGAGCATGGCTCTGAGCATGGCTCTTCCAAAGCCTTGTGACAACCAATTCCC 361  
 Db 61 CCAGAGATGGCTCTGAGCATGGCTCTTCCAAAGCCTTGTGACAACCAATTCCC 120  
 Qy 362 TCGATGATGATGCTCTGAGCTCTGCTGAGGAACTATGGAAAGCTCTGCCAGCAAGA 421  
 Db 121 TCGATGATGATGCTCTGAGCTCTGCTGAGGAACTATGGAAAGCTCTGCCAGCAAGA 180  
 Qy 422 AAATCTCTGCCAAAGCCCAAGCTTCTGTCACAGCTTCTGTCAGCTTCTGTCAGCTG 481  
 Db 181 AAATCTCTGCCAAAGCCCAAGCTTCTGTCACAGCTTCTGTCAGCTTCTGTCAGCTG 240  
 Qy 482 GAGGAGAGAGAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCG 541  
 Db 241 GAGGAGAGAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCG 300  
 Qy 542 GCGGAGCTGCTGGCTGAGCTGGGGAGCTGGCTCTGAGGATGGCTCTCTGAG 589  
 Db 301 GCGGAGCTGCTGGCTGAGCTGGGGAGCTGGCTCTGAGGATGGCTCTCTGAG 348

RESULT 5

US-09-867-550-1915  
 Sequence 1915; Application US/09867550  
 Patent No. US200408226A1  
 GENERAL INFORMATION:  
 / APPLICANT: Leach, Martin D.  
 / APPLICANT: Mehraban, Fuaad.  
 / APPLICANT: Conley, Pamela  
 / APPLICANT: Law, Debbie  
 / APPLICANT: Topper, James  
 / TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
 / FILE REFERENCE: 21402-013 (Curia-313)  
 / CURRENT FILING DATE: 2001-09-20  
 / PRIORITY APPLICATION NUMBER: USSN 60/208,427  
 / PRIOR FILING DATE: 2000-05-30  
 / NUMBER OF SEQ ID NOS: 2125  
 / SOFTWARE: FASTSEQ for Windows Version 4.0  
 / SEQ ID NO: 1915  
 / LENGTH: 875  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE: misc\_feature  
 / NAME/KEY: (1)  
 / LOCATION: (1)  
 / OTHER INFORMATION: Wherein n is one of a or t or c or g  
 / US - 09-867-550-1915

Query Match Similarity 28.8%; Score 341; DB 9; Length 875;  
 Best Local Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	Db	Score	DB	Length
843	2	GGATCCACTGCCTGACAAATGGCTGAGCTTACATCTCACCGGGCTCACCTTCCTCTCAC	902	61
Qy	903	TCCAGGCCCTGGGACCAATTACTCTGAGTGGGGATGACATCTGTGCTACTCAGG	962	62
Db	62	TCCAGGCCCTGGGACCAATTACTCTGAGTGGGGATGACATCTGTGCTACTCAGG	121	
Qy	963	AGCCCTGTCTGAGGATAACCTTACCTTGTA	1022	
Db	122	AGCCCTGTCTGAGGATAACCTTACCTTGTA	181	
Qy	1023	CTGTGAGAGGAAACCACTAACCTGAAAGAGCTGACACTCCCTCTGTTCTGAG	1082	
Db	182	CTGTGAGAGAACCATRACTGAGAAAGCTGACACTCCCTCTGTTCTGAG	241	
Qy	1083	CTGCCACAGGGGAGGAGTCCTCTCACTGAGGCTCTGGAGACTCCCTCTACA	1142	
Db	242	CTGCCACAGGGGAGGAGTCCTCTCACTGAGGCTCTGGAGACTCCCTCTACA	301	
Qy	1143	TCAGCCCTGAAGACAGGGCTCTTGGATGATGCCCTAG	1183	
Db	302	TCAGCCCTGAATGACAGGGCTCTTGGATGATGCCCTAG	342	

Query Match Similarity 11.3%; Score 134; DB 9; Length 432;  
 Best Local Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6  
 US-09-864-761-2829  
 Sequence 2829; Application US/09864761  
 / GENERAL INFORMATION:  
 / APPLICANT: Penn, Sharron G.  
 / APPLICANT: Rank, David R.  
 / APPLICANT: Hanzel, David K.  
 / APPLICANT: Chen, WenHeng  
 / TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 / GENE EXPRESSION ANALYSIS BY MICROARRAY  
 / FILE REFERENCE: Aeomica-X-1  
 / CURRENT APPLICATION NUMBER: US 09/864,761  
 / CURRENT FILING DATE: 2001-05-23  
 / PRIOR FILING DATE: 2000-02-04

RESULT 7  
 US-09-864-761-15513

Sequence 15513, Application US/09864761  
 Patent No. US20020045763A1  
 GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.  
 APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Weisheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEAR ACID PROBES USEFUL FOR

FILE REFERENCE: Aeonica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

SEQ ID NO: 15513

LENGTH: 448

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AL031662.24

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1

US-09-864-761-15513

Db 406 AAAGGCTGGACAG 419

RESULT 8

US-09-864-761-19612

Sequence 19612, Application US/09864761

Patent No. US20020045763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Weisheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEAR ACID PROBES USEFUL FOR

FILE REFERENCE: Aeonica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00660

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00659

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00658

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00657

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00656

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00655

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00654

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00653

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00652

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00651

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00650

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00649

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00648

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00647

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00646

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00645

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00644

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00643

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00642

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00641

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00640

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00639

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00638

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00637

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00636

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00635

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00634

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00633

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00632

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00631

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00630

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00629

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

LENGTH: 96

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AL031662.24

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1

SEQ ID NO: 19612

LENGTH: 96

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AL031662.24

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7e-01

OTHER INFORMATION: EST\_HUMAN HIT: A1125308.1, EVALUE 2.10e-01  
US-09-864-761-19612

Query Match Similarity 8.1%; Score 96; DB 9; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.3e-38;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 967 CTGTGTCCTGCAGGGCTGGCCGCTCCCTGCCAAGGADATACCCCTACTCTGACTGT 1026  
Db 1 CTGTGTCCTGCAGGGCTGGCCGCTCCCTGCCAAGGADATACCCCTACTCTGACTGT 60

Qy 1027 GGAGGACACCACTCACTGAAAGAGCTGACAG 1062  
Db 61 GGAGGACACCACTCAACTGAAAGAGCTGACAG 96

RESULT 9

US-09-814-153-17314

/ Sequence 17314, Application US/09814353  
/ Publication No. US20030165831A1

GENERAL INFORMATION:

/ APPLICANT: Lee, John

/ APPLICANT: Thompson, Pamela

/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND

/ TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

/ FILE REFERENCE: MRI-006B

/ CURRENT APPLICATION NUMBER: US/09/814,353

/ PRIOR APPLICATION NUMBER: 2001-03-21

/ PRIOR FILING DATE: 2000-03-21

/ PRIOR APPLICATION NUMBER: US 60/191,031

/ PRIOR FILING DATE: 2000-05-25

/ PRIOR APPLICATION NUMBER: US 60/211,940

/ PRIOR FILING DATE: 2000-06-15

/ PRIOR APPLICATION NUMBER: US 60/216,820

/ PRIOR FILING DATE: 2000-07-07

/ PRIOR APPLICATION NUMBER: US 60/220,661

/ PRIOR FILING DATE: 2000-07-25

/ PRIOR APPLICATION NUMBER: US 60/257,672

/ NUMBER OF SEQ ID NOS: 22037

/ SOFTWARE: FastSEQ for Windows Version 4.0

/ LENGTH: 320

/ ORGANISM: Homo sapiens

US-09-814-153-17314

Query Match Similarity 100.0%; Score 87; DB 10; Length 320;  
Best Local Similarity 100.0%; Pred. No. 7.6e-34; Mismatches 0; Indels 0; Gaps 0;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 549 TGTGCGTGAAGCTGGGAGGCAATTGACCACTGCTGAGGATGGAGACTGGTGACGG 608  
Db 103 TGTGCGTGAAGCTGGGAGGCAATTGACCACTGCTGAGGATGGAGACTGGTGACGG 162

Qy 609 TGTGCGTGAAGCTGGGAGGCAATTGACCACTGCTGAGGATGGAGACTGGTGACGG 635  
Db 163 TGTGCGTGAAGCTGGGAGGCAATTGACCACTGCTGAGGATGGAGACTGGTGACGG 189

RESULT 10

US-09-814-153-4631

/ Sequence 4631, Application US/09814353

/ Publication No. US20030165831A1

/ GENERAL INFORMATION:

/ APPLICANT: Lee, John

/ APPLICANT: Thompson, Pamela

/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND

/ ORGANISM: Homo sapiens



TITLE OF INVENTION: Polymorphisms in the Human Genome  
 FILE REFERENCE: 108827.129  
 CURRENT APPLICATION NUMBER: US/10/027,632  
 PRIOR APPLICATION NUMBER: US 60/104,302  
 PRIOR FILING DATE: 2002-04-30  
 PRIOR APPLICATION NUMBER: US 60/218,006  
 PRIOR FILING DATE: 2000-07-12  
 PRIOR APPLICATION NUMBER: US 60/198,676  
 PRIOR FILING DATE: 2000-04-20  
 PRIOR APPLICATION NUMBER: US 60/193,483  
 PRIOR FILING DATE: 2000-03-19  
 PRIOR APPLICATION NUMBER: US 60/185,218  
 PRIOR FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/167,363  
 PRIOR FILING DATE: 1999-11-23  
 PRIOR APPLICATION NUMBER: US 60/156,358  
 PRIOR FILING DATE: 1999-09-28  
 PRIOR APPLICATION NUMBER: US 60/146,002  
 PRIOR FILING DATE: 1999-08-09  
 NUMBER OF SEQ ID NOS: 325720  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 142058  
 LENGTH: 672  
 TYPE: DNA  
 ORGANISM: Human  
 US-10-027-632-142058

```

Query Match          1.7%; Score 20; DB 15; Length 672;
Best Local Similarity      100.0%; Pred. No. 9.5;
Matches   20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      747 CCTTCCTCATCGGGAGAC 766
Db      71 CCTTCCTCATCGGGAGAC 90
  
```

Search completed: February 20, 2004, 09:27:51  
 Job time : 487 secs





PF 10-JAN-2002; 2002WO-US000718.  
 PR 10-JAN-2001; 2001US-0260953P.  
 XX  
 PA (RIGE-) RIGEL PHARM INC.  
 XX  
 PI Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y, Sheng N;  
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC;  
 XX  
 DR WPI: 2002-575432/61.  
 DR N-PSDB; ABQ/4343.  
 XX  
 PT New src-like inhibitory molecule protein, useful for treating immunodeficiency disorders and inflammatory disorders, comprises N-terminal myristylation sequence, SH2 domain and/or SH3 domain.  
 PT  
 XX  
 PS Claim 7; FIG 2A; 91pp; English.  
 XX  
 CC The present sequence represents the human Src-like inhibitory molecule (SLIM) protein (I). The present invention describes a SLIM protein comprising an N-terminal myristylation sequence, an N-terminal SH2 domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising an N-terminal myristylation sequence and an N-terminal SH2 domain which is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and anti-HIV activities, and can be used as a modulator of lymphocyte activation, and of ubiquitination of a Cbl target protein, and in gene therapy. (I) is useful for screening a bioactive agent capable of binding to SLIM. (I) is also useful for screening a bioactive agent capable of modulating SLIM binding. (I) or its fragments is useful in the study or dysregulation of SLIM protein activity, i.e. to diagnose, treat or prevent SLIM associated disorders. (I) or the polynucleotide encoding it (II) is useful for modulating leukocyte and/or platelet activation, for modulating antigen receptor-induced signalling and activation in leukocyte and/or platelets and for modulating antigen receptor-induced signalling and activation in lymphocytes and/or mast cells. (I) or (II) is also useful for modulating the basal activity of lymphocytes. (I) or (II) is useful in the treatment of immunodeficiency disorders, such as acquired immunodeficiency syndrome (AIDS), for the prevention and treatment of acute inflammatory disorders, chronic inflammatory disorders, autoimmune disorder and transplant rejection  
 XX  
 SQ Sequence 261 AA;  
 Query Match 99.6%; Score 1347; DB 5; Length 261;  
 Best Local Similarity 99.6%; Fred. No. 6.8e-130;  
 Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 DR WJ; 2002-463632/49.  
 DR N-PSDB; AAD43980.  
 XX  
 QY 1 MGSLPSRRKSPLSPSPLSSVQGQPTMMEARSKATAVALGSFPAGGPABSLRLGEPLT 60  
 Ds 1 MGSLPSRRKSPLSPSPLSSVQGQPTMMEARSKATAVALGSFPAGGPABSLRLGEPLT 60  
 QY 61 IYSEDGDWWTTLSEVSGREYNIPSTHGKSHGNLYEGISPEKAEEELLIPNGNGGATL 120  
 Ds 61 IYSEDGDWWTTLSEVSGREYNIPSTHGKSHGNLYEGISPEKAEEELLIPNGNGGATL 120  
 CC The invention relates to a substantially purified human SH2/SN3-domain-containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-2). The invention is useful for treating an immune disorder involving hyperactivation of B- or T-lymphocytes in a mammal. The invention is useful for screening for antagonists or inhibitors of the interaction of hSLAP-2 with cellular signalling compounds, for diagnosing, treating or preventing diseases or disorders associated with aberrant or uncontrolled cellular signal transduction, for determining those cellular signalling molecules which associate with hSLAP-2 and which provide critical signals for cell activation, and as effectors in methods to affect T-cell activation. The invention is useful in screening assays to identify and detect candidate biactive agents that modulate hSLAP-2 biactivity, for potential use to treat autoimmune diseases which may be caused by hyperactivated B cells, as well as to treat diseases which may be caused by hyperactivated T cells, in addition to other immune system related conditions, diseases, or disorders. T-cell and B-cell neoplasms, inflammation disorders, diseases and conditions, rheumatoid arthritis, osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's  
 CC RESULT 4  
 ID AAE26357 standard; Protein: 261 AA.  
 XX

Page 4

Sequence 261 AA:					
Query	Match	Score	Length	DB	
99.6%;	Score 1347;	9.6%;	Length 261;		
Best Local Similarity	Pred. No.	6.8e-130;			
Matches 260;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0
1 MGSLPSRKSLPSLSSVQGQPYTMEAERSKATAVALGSFPAGGPABLSRLRGEPLT	60				
1 MGSLPSRKSLPSLSSVQGQPYTMEAERSKATAVALGSFPAGGPABLSRLRGEPLT	60				
61 IYEDGMWTVLSEVSREYNIPSYVKGKVSHGWLYEGLSREKABELLLIPGNNGGAFL	120				
61 IYEDGMWTVLSEVSREYNIPSYVAKVSHGWLYEGLREKABELLLIPGNNGGAFL	120				
121 RESQTGRGSSYSLSRPASWDRIRHYRHCLDNGLWLYISPRLEPPSLOALVTHYSELA	180				
121 RESQTGRGSSYSLSRPASWDRIRHYRHCLDNGLWLYISPRLEPPSLOALVTHYSELA	180				
181 DDDCCLIKEPCVYIQRAGPLPGKDIDPLPVTVORTPLNKELDSSLIFSEATGBSSLSEG	240				
181 DDDCCLIKEPCVYIQRAGPLPGKDIDPLPVTVORTPLNKELDSSLIFSEATGBSSLSEG	240				
241 LRESLSFVISLNDAEVSLDDA	261				
241 LRESLSFVISLNDAEVSLDDA	261				

RESULT 6	AAO15456			
	ID AAO15456	standard; protein; 259 AA.		
	XX			
	AC AAO15456;			
			03-OCT-2002 (first entry)	
Homo sapiens.				
WO2000058473-A2.				
05-OCT-2000.				
31-MAR-2000; 2000WO-US008621.				
31-MAR-1999; 99US-0127607P.				
02-APR-1999; 99US-0127735P.				
05-APR-1999; 99US-0127728P.				
30-MAR-2000; 2000US-00540763.				

KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder; immunosuppression; myeloproliferative disorder; breast cancer.

XX Mus sp.

OS WO200242452-A2.

XX PD 30-MAY-2002.

XX PP 26-NOV-2001; 2001WO-CA001662.

XX PR 27-NOV-2000; 2000CA-02324663.

XX PA (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX PI McGlade JC, Loreto MP;

XX DR WPI; 2002-566564/60.

XX PT DR-N-PSDB; AAL44087.

XX PT New isolated modulator of antigen receptor signalling protein or its fragment, useful for treating malignant disorders such as myeloid malignancies, autoimmune disorders and myeloproliferative disorders.

XX PS Claim 7; Fig 1A; 110pp; English.

CC The invention comprises the amino acid and coding sequences of modulator of antigen receptor signalling (MARS) proteins. The MARS protein is a purative tumour suppressor gene and exhibits structural and sequence similarity to the Scr-like adaptor protein (SlAP). The MARS DNA and protein sequences of the invention are useful for the treatment of myeloid malignancies (e.g. acute myelogenous leukaemia), autoimmune disorders, immunosuppression, myeloproliferative disorders and malignancies related to the de-regulation of tyrosine kinases (e.g. breast cancer). The present amino acid sequence represents a mouse MARS protein.

SQ Sequence 259 AA;

Query Match 76.0%; Score 1028; DB 5; Length 259;  
Best Local Similarity 79.4%; Pred. No. 4.e-97;  
Matches 208; Conservative 16; Mismatches 34; Indels 4; Gaps 3;

Qy 1 MGSLPSRKSPLSPSLSVQGQPVTNEAERSKATAVALGSFPAGGAELSLRLGEPLT 60  
Db 1 MGSLSSRGKT-SSPSPSGGDQEPMQPERHKVATAVGSPAGEQARSLRLGEPLT 59

Qy 61 IVSEGDWWTVLSEVSGREYNIPSPVHYCKVSHGWLYEGLSRKEAELLPGNPGGAFLI 120  
Db 60 ISEGDWWTVQSEVSGREYHMSVYAVKAWHGWLYEGLSRKEAEELLPGNPGGAFLI 119

Qy 121 RESQTRGYSYSLSVRLSPASWRIRHYRTHICLDNGNLYISPRLTPESLQALVDHYSLEA 180  
Db 120 RESQTRGYSYSLSVRLSPASWRIRHYRTHICLDNGNLYISPRLTPESLQALVDHYSLEA 179

Qy 181 DDICCLLKEPVCUQAGPLPKDQPLPYTVQRTPLNWKELDSLSEA-ATGEESLSE 239  
Db 180 DGICCLLKEPVCUQAGPLPKDQPLPYTVQRTPLNWKELDSLLEAPSGEASLSE 239

Qy 240 GLRSLSFPISLNDEAVSLDDA 261  
Db 240 GLRSLSSTISLAAED--PLDDA 259

Query Match 76.0%; Score 1028; DB 5; Length 259;  
Best Local Similarity 79.4%; Pred. No. 4.e-97;  
Matches 208; Conservative 16; Mismatches 34; Indels 4; Gaps 3;

Qy 1 MGSLPSRKSPLSPSLSVQGQPVTNEAERSKATAVALGSFPAGGAELSLRLGEPLT 60  
Db 1 MGSLPSRKSPLSPSLSVQGQPVTNEAERSKATAVALGSFPAGGAELSLRLGEPLT 60

Qy 61 IVSEGDWWTVLSEVSGREYNIPSPVHYCKVSHGWLYEGLSRKEAELLPGNPGGAFLI 120  
Db 60 ISEGDWWTVQSEVSGREYHMSVYAVKAWHGWLYEGLSRKEAEELLPGNPGGAFLI 119

Qy 121 RESQTRGYSYSLSVRLSPASWRIRHYRTHICLDNGNLYISPRLTPESLQALVDHYSLEA 180  
Db 120 RESQTRGYSYSLSVRLSPASWRIRHYRTHICLDNGNLYISPRLTPESLQALVDHYSLEA 179

Qy 181 RESQTRGYSYSLSVRLSPASWRIRHYRTHICLDNGNLYISPRLTPESLQALVDHYSLEA 178  
Db 180 RESQTRGYSYSLSVRLSPASWRIRHYRTHICLDNGNLYISPRLTPESLQALVDHYSLEA 180

Qy 179 ----LADDICCLLKEPVCUQAGPLPKDQPLPYTVQRTPLNWKELDSLSEA-ATGEESLSE 239  
Db 181 PAPWQGYTPTCDCDAEDTQLERACQLP 200

RESULT 8  
ABP6788  
ID ABP6788 standard; protein; 210 AA.  
XX AC ABP6788;  
XX DT 25-FEB-2003 (first entry)  
XX DE Human protein SEQ ID 448.

Human; expressed sequence tag; EST; haematopoietic disorder;  
 KW central nervous system disease; viral infection;  
 KW peripheral nervous system disease; nonhealing wound; infectious disease;  
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
 KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;  
 KW anti-allergic; antiinflammatory; immunosuppressive; neuroprotective;  
 KW cytostatic; haemostatic; viricide; antibacterial; fungicide;  
 KW immunostimulant; cerebroprotective;  
 KW immunoprotective; fungicide;

XX  
 Homo sapiens.XX  
 WO200259260-A2.XX  
 PN 01-AUG-2002.XX  
 PD 16-NOV-2001; 2001WO-US042936.XX  
 PR 17-NOV-2000; 2000US-00714936.XX  
 PA (HYSEQ INC.XX  
 PI Tang YR, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;

PI Ren F, Xue AJ, Yang Y, Weinman T, Brmanac RT;

XX  
 DR 2002-590824/63.

N-PSDB; ABQ99374.

XX  
 PT New isolated polynucleotide, useful in research, diagnostic or

PT therapeutic methods, e.g. preventing or treating disorders involving

PT aberrant protein expression or biological activity.

XX  
 XX  
 PS Claim 20; SEQ ID NO 448; 394pp; English.XX  
 CC The present invention relates to novel human coding sequences (ABQ99268-

CC ABQ99308) and proteins (ABP64182-ABP6502). The sequences are useful in

CC therapeutic, diagnostic and research methods. The polynucleotides may be

CC used in the field of molecular biology as hybridisation probes, primers

CC for PCR, for chromosome and gene mapping, for the recombinant production

CC of protein, or in generation of anti-sense DNA or RNA. The

CC polynucleotides are useful in diagnostics as expressed sequence tags

CC (ESTs) for identifying expressed genes or for physical mapping of the

CC human genome. The proteins may be used as molecular weight markers, as

CC nutritional sources or supplements. The proteins may be used to maintain

CC and expand cell population in a totipotent or pluripotent state

CC useful for re-engineering damaged or diseased tissues, transplantation,

CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The

CC polynucleotides and proteins are useful for preventing, treating or

CC ameliorating disorders involving aberrant protein expression or

CC biological activity, e.g. haematopoietic disorders, central/peripheral

CC nervous system diseases, mechanical and traumatic disorders, non-healing

CC wounds, immune deficiencies and disorders, infectious diseases caused by

CC viral, bacterial or fungal infection, autoimmune disorders, allergic

CC reactions and conditions, coagulation disorders, or cancer. The

CC polynucleotide sequences of the invention were assembled from ESTs

CC isolated mainly by sequencing by hybridisation, and in some cases,

CC sequences obtained from one or more public databases. Note: The sequence

CC data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from WIPO at

XX  
 ftp://wipo.int/pub/published\_pct\_sequencesXX  
 Sequence 210 AA;

SQ Query Match 68.3%; Score 923.5; DB 5; Length 210;

Best Local Similarity 87.9%; Pred. No. 2.1e-86;

Matches 182; Conservative 3; Mismatches

Indels 7; Gaps 1;

XX  
 Sq Query 1 MGSLPSRKSLPSLSSVQGPVMEARSKATAVALGSFAGGPAAELSLRGEPT 60

Db 1 MGGLPSRKSLPSLSSVQGPVMEARSKATAVALGSFAGGPAAELSLRGEPT 60

XX  
 Sq Query 61 IVSEDGDWWTIVLSEVGREYNTPSVHGKVSHGMVYEGISREKARELLIILPGNPGAAFLI 120

Db 61 IVSEDGDWWTIVLSEVGREYNTPSVHAKVSHGMVYEGISREKARELLIILPGNPGAAFLI 120

Qy	121 RESQTGRGYSISVRLSPASWDR1PHYR1HC1DNGWL1SPLTTPSLOALVDHYSE--
Db	121 RESQTGRGYSISVRLSPASWDR1PHYR1HC1DNGWL1SPLTTPSLOALVDHYSE--
Qy	179 ----LADDICCLLKEPCVLRAGPQP 200
Db	181 PAPWQGYTPCTCDCAEDTTQLERAGQP 207
RESULT 9	
ID ABP64107	
XX	
AC ABP64107;	
XX	
DT 04-NOV-2002 (first entry)	
XX	
DE Human ORF477.	
XX	
Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;	
KW antiinflammatory; gene therapy; human; ORF8; atherosgenic; Platelet;	
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;	
KW cancer; cardiovascular disease; allergy; autoimmune disease;	
KW wound healing; blood coagulation disorder; inflammatory disorder.	
XX	
OS Homo sapiens.	
XX	
PN US2002082206-A1	
XX	
PR 2000-2000; 2000US-0208427P.	
XX	
PD 27-JUN-2002.	
XX	
PA 30-MAY-2001; 2001US-00867550.	
XX	
PR 30-MAY-2000; 2000US-0208427P.	
XX	
PA (LEAC) LEACH M. D.	
PA (MEHR) MEHRABAN F.	
PA (CONL) CONLEY P. B.	
PA (TOPP) TOPPER J. N.	
(LAWD) LAW D.	
XX	
Leach MD, Mehrabian F, Conley PB, Topper JN, Law D;	
XX	
DR WBI: 2002-626554/67.	
DR N-PSDB; ABQ98670.	
XX	
New polypeptide designated ORFX are present in human atherogenic cells	
PR and are useful to prevent and treat ORFX-associated disorders including	
cancer, allergy, wound healing or autoimmune, cardiovascular or	
inflammatory disease.	
XX	
Claim 10; SEQ ID NO 954; 78pp; English.	
XX	
The present invention relates to novel human ORFX Polypeptides and their	
CC coding sequences (ABP63631-ABP64581 and ABQ98194-ABQ92267). The sequences	
CC were discovered in human atherogenic cells, in particular in platelets	
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in	
CC many other tissues as well. Atherogenic cells are cells which have the	
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and	
CC nucleic acids are useful for treating or preventing a pathological	
CC condition associated with an ORFX-associated disorder, e.g. cancer,	
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood	
CC coagulation disorders or inflammatory disorders. Note: The sequence data	
CC for this patent did not form part of the printed specification, but was	
CC obtained in electronic format directly from the USPRO web site at	
CC seqdata.uspto.gov/sequence.html?DocID=20020082206	
XX	
SQ Sequence 159 AA;	
Query Match 61.0%; Score 826; DB 5; Length 159;	
Best Local Similarity 99.4%; Pred. No. 1.5e-76;	
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

Qy 1 MGSLSRSRKSLPSLSSVQGQGPVMEARSKATAVALGSFPAGGPALSIRIGEPLT 60  
 Db 1 MGSLSRSRKSLPSLSSVQGQGPVMEARSKATAVALGSFPAGGPALSIRIGEPLT 60  
 Qy 61 IVSEGDGVWTVLSESVGREYNIIPSVHVGKVSHGWLYEGLSREKAELLLPENPGGAFLI 120  
 Db 61 IVSEGDGVWTVLSESVGREYNIIPSVHVGKVSHGWLYEGLSREKAELLLPENPGGAFLI 120  
 Qy 121 RESQTTRGSYSSLSTRLSPASWRIRHYRICHLDNGWLY 159  
 Db 121 RESQTTRGSYSSLSTRLSPASWRIRHYRICHLDNGWLY 159

## RESULT 10

ABP64588 ID ABP64588 standard; protein; 113 AA.  
 AC ABP64588;  
 XX  
 DT 04-NOV-2002 (first entry)  
 XX

Human ORF958.

KW Cystostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antiinflammatory; Gene therapy; human; QEFX; atherogenic; Platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 cancer; cardiovascular disease; allergy; autoimmune disorder;  
 wound healing; blood coagulation disorder; inflammatory disorder.

Homo sapiens.

US2002082206-A1.

XX

PD 27-JUN-2002.

XX

PF 30-MAY-2001; 2001US-00867550.

XX

PR 30-MAY-2000; 2000US-0208427P.

XX

(LEAAC/)

LEACH M D.

PA (MEHR/)

MEHRABAN P.

PA (CONL/)

CONLEY P B.

PA (TOPP/)

TOPPER J N.

PA (LAWD/)

LAW D.

XX

PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D.

XX

DR WPI; 2002-626554/67.

DR N-PSDB; ABQ9151.

XX

PT New polypeptide designated ORFX are present in human atherosgenic cells and are useful to prevent and treat ORFX-associated disorders including

cancer, allergy, wound healing or autoimmune, cardiovascular or

inflammatory disease.

XX

Claim 10; SEQ ID NO 1916; 78pp; English.

DR N-PSDB; ABQ9151.

XX

PT New polypeptide designated ORFX are present in human atherosgenic cells and are useful to prevent and treat ORFX-associated disorders including

cancer, allergy, wound healing, blood coagulation disorders, autoimmunity disease, wound healing, blood coagulation disorders or inflammatory disorders. Note: The sequence data obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov/sequence.html?DocID=20020082206

Sequence 113 AA;

Query Match 43.3%; Score 586; DB 5; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-52;  
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 150 IHCNDGWLWISPRUTPSIQLAVDHYSELADDICCLKEPCVLRQAGPLPGKDIPLYT 209  
 Db 2 IHCNDGWLWISPRUTPSIQLAVDHYSELADDICCLKEPCVLRQAGPLPGKDIPLYT 61  
 Qy 210 VQRTPLNKELDSLIFSEATGEBSLLSEGREGESLLSEGREGESLLSEGREGESLL 261  
 Db 62 VQRTPLNKELDSLIFSEATGEBSLLSEGREGESLLSEGREGESLLSEGREGESLL 113

## RESULT 11

ABP52188 ID ABP52188 standard; protein; 276 AA.

AC ABP52188;

XX

DT 15-OCT-2002 (first entry)

XX

Human Src-like adapter protein (SLAP).

DE DE

XX

KW Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;

KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;

KW modulator; lymphocyte; gene therapy; immunodeficiency disorder;

KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;

KW chronic inflammatory disorder; autoimmune disorder; transplant rejection.

XX

OS Homo sapiens.

OS

XX

PN WO200255707-A2.

XX

PD 18-JUL-2002.

XX

PF 10-JAN-2002; 2002WC-US000718.

XX

PR 10-JAN-2001; 2001US-02609531P.

XX

(RIGE-) RIGEL PHARM INC.

PA PA

XX

PA Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;

PI Payan DG, Mancebo HS, Wu J, Zhou X, Sheng N;

XX

WPI; 2002-57432/61.

XX

PT New src-like inhibitory molecule Protein, useful for treating

PT immunodeficiency disorders and inflammatory disorders, comprises N-

PT terminal myristylation sequence, SH2 domain and/or SH3 domain.

XX

PS Example 2; FIG 2B; 91pp; English.

XX

CC The present invention describes the human Src-like inhibitory molecule

CC (SLIM) protein (I). The present invention describes a SLIM protein

CC comprising an N-terminal myristylation sequence, an N-terminal SH2

CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising

CC an N-terminal myristylation sequence and an N-terminal SH2 domain which

CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and

CC anti-HIV activities, and can be used as a modulator of lymphocyte

CC activation, and of ubiquitination of a Cbl target protein, and in gene

CC therapy. (I) is useful for screening a bioactive agent capable of binding

CC to SLIM. (I) is also useful for screening a bioactive agent capable of

CC modulating SLIM binding. (I) or its fragments is useful in the study or

CC in the treatment of conditions which involves this function or

CC deregulation of SLIM protein activity, i.e. to diagnose, treat or

CC prevent SLIM associated disorders. (I) or the polynucleotide encoding it

CC (I) is useful for modulating leukocyte and/or platelet activation, for

CC modulating antigen receptor-induced signalling and activation in

CC leukocyte and/or platelets and for modulating antigen receptor-induced

CC signalling and activation in lymphocytes and/or mast cells. (I) or

CC (II) is useful for modulating the basal activity of lymphocytes, (I) or

CC (II) is useful in the treatment of immunodeficiency disorders, such as



Db 230 RESTASYNLSLTSE 242  
 RESULT 13  
 ABR59698 standard; protein; 276 AA.  
 ID ABR59698;  
 AC ABR59698;  
 XX DT 25-JUL-2003 (first entry)  
 XX Human Src-like-adaptor.  
 DE Human; T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; asthma;  
 KW immunosuppressive; antiasthmatic; antiallergic; antiinflammatory;  
 KW lymphocyte activation; lymphocyte migration; cytokine production;  
 KW cell surface marker expression; antibody production; apoptosis; allergy;  
 KW antibody proliferation; antibody differentiation; hypersensitivity;  
 KW graft versus host disease; inflammation; Src-like-adaptor.  
 XX OS Homo sapiens.  
 XX PN WO2003029277-A2.  
 XX PD 10-APR-2003.  
 XX PF 02-OCT-2002; 2002WO-US031618.  
 XX PR 03-OCT-2001; 2001US-0327212P.  
 XX PA (RIGE-) RIGEL PHARM INC.  
 XX PI Chu P, Li C, Liao XC, Masuda E, Pardo J, Zhao H;  
 XX DR WPI; 2003-363276/34.  
 XX N-PSDB; ACCB1091.  
 PT Identifying a compound that modulates T lymphocyte activation, useful for  
 PT monitoring changes in cell surface marker expression, comprises  
 PT contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with  
 PT a compound.  
 XX Disclosure; Page 72; 126PP; English.  
 XX The invention relates to a novel method for identifying a compound that  
 CC modulates T lymphocyte activation. The method comprises contacting a T  
 CC cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,  
 CC where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic  
 CC acid that hybridises to a nucleic acid encoding a polypeptide having a  
 CC sequence selected from two 506-amino acid sequence and a 415-amino acid  
 CC sequence given in the specification. The method of the invention has  
 CC immunosuppressive, antiasthmatic, antiallergic, and antiinflammatory  
 CC activity. The method is useful for identifying compounds that modulate  
 CC lymphocyte activation and migration, and for monitoring changes in cell  
 CC surface marker expression, cytokine production, antibody production,  
 CC proliferation and differentiation, and apoptosis, using either cell lines  
 CC or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as  
 CC drug targets for compounds that suppress or activate lymphocyte  
 CC activation and migration, e.g. for the treatment of diseases in which  
 CC modulation of the immune response is desired such as delayed type  
 CC hypersensitivity reactions, asthma, allergies, graft versus host disease,  
 CC and acute and chronic inflammation. Modulators of lymphocyte activation  
 CC are useful for treating disorders related T and B cell activation and  
 CC migration. The present sequence is used in the exemplification of the  
 CC invention.  
 XX Sequence 276 AA;  
 SQ Query Match 35.6%; Score 481.5; DB 6; Length 276;  
 Best Local Similarity 40.3%; Prd. No. 1.1e-40;  
 Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;

Db 6 KSTPAPA-----ERPLPNPEGLDSDPLAVLSDYPSSEDISPPIFRRCGKLRVISDEGGW 58  
 QY 69 WTVLSEVSREYNTPSVHGVSHGWLTGGLSLSEKABILLLGNGGAFLINESQTRG 128  
 Db 59 WKAISLTSLTSRRESYIPGICVARYHGWLPGDGRDKABELLQLPDTRKGSFMRESETKG 118  
 QY 129 SYSLSVRLSRPASMDRIRHYTHCLDNQWLYSPRLTSPSLOALVDHYSELADDICCLK 188  
 Db 119 FYSLSVR----HRQVHYRIPLPNWMYXISPRLTQCLBLVNHYSEVAIGLCCLT 172  
 QY 189 EPCVILQRAGPLPGKDIPLPVQTPLNWKDSSLIFSEATG-----BESLISFGI 241  
 Db 173 TPCUQTSTAPAYRASSSPVTLQRKTWDRRSR---IQEDPEGTENPLGVDELSFSYGL 229  
 QY 242 RESTSFYISINDE 254  
 Db 230 RESTASYNLSLTSE 242

RESULT 14  
 AAU1072  
 ID AAU1072 standard; protein; 315 AA.  
 XX AAU1072;  
 AC AAU1072;  
 XX DT 18-DEC-2001 (first entry)  
 XX Novel human secreted protein #1563.  
 DE Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematoopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory;  
 XX Homo sapiens.  
 OS WO200179449-A2.  
 XX PN 2001079449-A2.  
 XX PD 25-OCT-2001.  
 XX PR 16-APR-2001; 2001WO-US008656.  
 XX PR 18-APR-2000; 2000US-00552929.  
 XX PR 26-JAN-2001; 2000US-00770160.  
 XX (HYSEQ INC.  
 PA Tang YT, Liu C, Drmanac RT;  
 DR WPI; 2001-611725/70.

Nucleic acids encoding a range of human polypeptides, useful in genetic vaccination, testing and therapy.  
 XX Claim 20; Page 399; 765PP; English.  
 XX The invention relates to novel human secreted polypeptides. The proteins  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in general vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haemopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU9510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention.

Qy 9 KSLPSPSLSSVQGQGPVTTMEAERSKATAVALGSFPAGGPAAELSRLGEPTIVSSEDGW 68

XX Sequence 315 AA;  
 SQ Score 470.5; DB 4; Length 315;  
 Query Match Best Local Similarity 39.3%; Pred. No. 1.8e-39;  
 Matches 103; Conservative 46; Mismatches 100; Indels 13; Gaps 4;  
 PSSRKSTPSLSSVQGQPYTMEEAERSKATAVALGSFPAGGPAAELSLRIGEPLTIVSE 64  
 |:|:|:|:|:|:|:|:  
 53 GPKKENGNMSMSTPAPAEPRLPNPEQLDSDFLAVISDYPFDISPIFRGEKLVISD 92  
 DGDWWTVLSEVSGREYNIPSPVHGVYSHG-WIYEGLSREKAELLIPGNGGAFLIRES 123  
 QY 65 :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 93 EGGRWKAISLSTSRESTIPGICVARYHGLWLFEGLGRDAEELLQLPDTKVGSMIRES 152  
 XX  
 QY 124 QPRRGYSLSISVLSSRPSWDRPRHRTRHCLNGWLNISPRITFPLQALWDHYSELADDI 183  
 Db 153 ETIKGFYSLSVR----HRQVKTYR-FRLPNWNYTISPRLTFOCLEDLYHSEVAQDI 206  
 CCILKEPVULQRAQPLPOKDIPILPVTWQRTPLNWKELDISLJLFSEATGE---ESLIS 238  
 QY 184 CCVLTTPCLQTSTAAPAVRACSPPTVLRKTYDWRAY-SRLQEDPGETENPLGTVESLIS 265  
 Db 207 EGIRELSLIFYISNDNEAVSLDD 260  
 QY 239 YGLRESTASYLTSLEDISSF D 287  
 Db 266 YYLRESTASYLTSLEDISSF D 287  
 RESULT 15  
 ABP64106 DT 04-NOV-2002 (First entry)  
 ID ABP64106 standard; protein; 96 AA.  
 XX AC ABP64106;  
 XX DT 04-NOV-2002 (First entry)  
 XX DB Human ORF476.  
 XX Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 Antinflammatory; Gene therapy; human; ORFX; atherogenic; platelet;  
 human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 cancer; cardiovascular disease; allergy; autoimmune disease;  
 wound healing; blood coagulation disorder; inflammatory disorder.  
 Homo sapiens.  
 PN US2002082206-A1.  
 XX 27-JUN-2002.  
 XX 30-MAY-2001; 2001US-00867550.  
 XX 30-MAY-2000; 2000US-0208427P.  
 XX (LEAC/) LEACH M. D.  
 PA (MEHR/) MEHRABAN F.  
 PA (CONL/) CONLEY P. B.  
 PA (TOPP/) TOPPER J. N.  
 PA (LAWD/) LAW D.  
 XX WPI; 2002-626554/67.  
 DR N-PSDB; ABQ98669.  
 XX New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 XX Claim 10; SEQ ID NO 952; 78pp; English.  
 XX The present invention relates to novel human ORFX polypeptides and their

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Inc.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:19:42 ; Search time 45 Seconds  
(without alignments)  
557.911 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLPSRRKSLPSLSSSV.....RESLSFVSLNDEAVSLLDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

PIR\_78:\*

1: Piri:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	493.5	36.5	281	2 A57152	src-like adaptor protein
2	370.5	27.4	512	1 TVHUY	protein-tyrosine
3	367.5	27.2	512	1 A39719	protein-tyrosine
4	367.5	27.2	512	1 TVHDFC	protein-tyrosine
5	360.5	26.6	505	1 JQ1321	protein-tyrosine
6	353.5	26.1	503	1 JQ1321	protein-tyrosine
7	352.5	26.1	505	2 TMNSHC	protein-tyrosine
8	352	26.0	503	1 TMNSHC	protein-tyrosine
9	344.5	25.5	499	1 A4092	protein-tyrosine
10	340	25.1	509	1 ORHULK	protein-tyrosine
11	338.5	25.0	509	1 I4810	protein-tyrosine
12	333	24.6	507	1 A39939	protein-tyrosine
13	325	24.0	539	1 B49114	protein-tyrosine
14	320	23.7	528	1 TVEFG9	protein-tyrosine
15	319	23.6	537	1 A45501	protein-tyrosine
16	317	23.4	541	1 TVCHYS	protein-tyrosine
17	315.5	23.3	537	1 TWHUSY	protein-tyrosine
18	315.5	23.3	543	1 TWHUSY	protein-tyrosine
19	312.5	23.1	529	1 TWDYFR	protein-tyrosine
20	311	23.0	544	2 I15193	protein-tyrosine
21	310.5	22.9	537	1 A41806	protein-tyrosine
22	309	22.8	534	1 A41991	protein-tyrosine
23	307	22.7	517	2 S24447	protein-tyrosine
24	306.5	22.7	541	2 S31645	protein-tyrosine
25	305.5	22.6	537	2 I15192	protein-tyrosine
26	301	22.2	517	2 A43107	protein-tyrosine
27	301	22.2	536	2 S33569	protein-tyrosine
28	297.5	22.0	534	1 A43114	protein-tyrosine
29	295.5	21.8	542	2 A43114	protein-tyrosine

## ALIGNMENTS

RESULT 1  
A57152  
src-like adaptor protein - mouse  
C Species: Mus musculus (house mouse)  
C; Date: 08-Dec-1995 #sequence\_revision 08-Dec-1995 #text\_change 12-Feb-1999  
C; Accession: A57152  
R; Pandey, A.; Duan, H.; Dixit, V. M.  
J. Biol. Chem. 270, 19201-19204, 1995  
A; Title: Characterization of a novel Src-like adapter protein that associates with the E<sub>c</sub>  
A; Reference number: A57152; PMID:7543898  
A; Accession: A57152  
A; Status: Preliminary; not compared with conceptual translation  
A; Molecule type: mRNA  
A; Residues: 1-281 <PAN>  
A; Cross-references: GB:U29056  
C; Superfamily: SH3 homology; SH2 homology  
F: 29-77/Domain: SH3 homology <SH3>  
F: 84-175/Domain: SH2 homology <SH2>

Query Match 36.5%; Score 493.5%; DB 2; Length 281;  
Best Local Similarity 43.6%; Prod. No. 2.1e-33; Mismatches 82; Indels 29; Gaps 7;

Matches 145; Conservative 38; Mismatches 82;

QY 9 KSLPSPS---LSSVQGGQPVTMAERSKATAYALGSPAGGPAAELSRGLGRPLTVISED 65  
Db 6 KSTPPPSRPLSSS-----EGLEDFFLAY-LTDPYSSD1SPPIFRGERLRLSDE 55

QY 66 GDWITVLSEVSGREYNINPSVAVGKVSHGWLYEGSRSKAEELLILPGNGAFLIRESQQT 125  
Db 56 GHWWKAISLSTGREGESYIGCIVARYVSGWLFEGLGRKAEELLQLPDKIGSMIRESET 115

QY 126 RRGYSLSYLRLSPASWRHRHICLDNGWLYISPLTFPSLQALVDHYSSLADDCC 185  
Db 116 KKGPYSLSYR-----HQVQHYRIFRPNNWYIISPLTFQCLEDLVTHYSEVADGCC 169

QY 186 LKEBPCVLR-----AGPLPGKDIPLPVTVORTPLAKWELDSSLLFSEBAATG ---EBSL 236  
Db 170 VLTPLCLAMNPAPTSHS-----BSPCTSPGSPTLQRQDNTRVSLQEGSEGARNPLRVDSL 229

QY 237 LSEGIRESLSFYISL-NDEAVSUD 259  
Db 230 PSYGLRESIASYSLTGDDSSSPD 253

RESULT 2  
TVHUY  
Protein-tyrosine kinase (EC 2.7.1.112) lym, splice form A - human  
N; Contains: protein-tyrosine kinase lym, splice form B  
C; Species: Homo sapiens (man)  
C; Date: 31-Mar-1999 #sequence\_revision 31-Mar-1999 #text\_change 18-Feb-2000  
C; Accession: A26719; D38268; E4949; I53715  
R; Yamanashi, Y.; Fukushima, S.I.; Semba, K.; Miyajima, N.; Matsubara, K.;

Mol. Cell. Biol. 7, 237-243, 1987	A;Title: The Yes-related cellular gene lyn encodes a possible tyrosine kinase similar to protein-tyrosine kinase (EC 2.7.1.112) lyn, splice form A - rat	RESULT 3
A;Reference number: A26719; MUID:87172710; PMID:3561390	A;Accession: A26719	I56160
A;Molecule type: mRNA	A;Status: not compared with conceptual translation	protein-tyrosine kinase (EC 2.7.1.112) lyn, splice form A - rat
A;Residues: 1-512 <RID>	A;Residues: 369-424 <PAR>	N;Contains: protein-tyrosine kinase (Norway rat)
A;Cross-references: GB:MI6038; NID:9187268; PIDN:AAA59540.1; PID:9307144	R;Bielek, W.; Ziemieli, A.; Kappos, L.; Miescher, G.C.	C;Species: Rattus norvegicus (Norway rat)
J. R. Partanen, J.; Maekela, T.P.; Alitalo, K.; Lehtesaloaho, H.; Alitalo, K.	Biochem. Biophys. Res. Commun. 186, 1403-1409, 1992	C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
A;Title: Expression of the B cell-associated tyrosine kinase gene lyn in primary neuroblastoma cells	A;Reference number: PH0949; MUID:92378604; PMID:1510669	R;Minoguchi, K.; Nishikata, H.; Siraganian, R.P.
A;Cross-references: PH0949	A;Accession: D38268	J. Immunol. 150, 222, 1993
A;Molecule type: mRNA	A;Accession: D38268	A;Title: Bacterially expressed rat p56lyn binds several proteins in rat basophilic leuke-
A;Residues: 369-424 <BEI>	A;Accession: I56160	A;Reference number: I56160
A;Experimental source: neuroblastoma SK-N-T cell	A;Status: Preliminary; translated from GB/EMBL/DDJB	A;Accession: I56160
R;Rider, L.G.; Raben, N.; Miller, L.; Jelisema, C.	A;Molecule type: mRNA	A;Status: Preliminary; translated from GB/EMBL/DDJB
A;Residues: 1-24-46-12 <RID>	A;Accession: I53715	A;Molecule type: mRNA
A;Cross-references: GB:MT9321; NID:9187270; PIDN:AAB50019.1; PID:9187271	A;Accession: I53715	A;Residues: 1-24,46-12,46-12 <RID>
A;Experimental source: splice form B	A;Accession: I53715	A;Cross-references: GB:114823; NID:9294580; PIDN:AAA20945.1; PMID:9294581
C;Genetics:	A;Accession: I53715	A;Note: in Genbank entry RATLYNATR, release 116 0, PIDN:AAA20944.1, the source is design-
A;Gene: GDP:LYN	A;Accession: I53715	A;Title: The cDNAs encoding two forms of the LYN protein tyrosine kinase are expressed in rat lymphocytes
A;Cross-references: GDB:120159; OMIM:165120	A;Accession: I53715	A;Reference number: I53715
C;Function:	A;Accession: I53715-8Qter	A;Accession: I67811
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP	A;Accession: I53715-8Qter	A;Accession: I67811
C;Superfamily: protein-tyrosine kinase src; Protein kinase homology: SH3 h	A;Accession: I53715-8Qter	A;Accession: I67812
C;Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprotein	A;Accession: I53715-8Qter	A;Accession: I67812/
Y;R;O;S;P;L;P;S;L;U;S;V;Q;G;G;P;T;M;E;R;S;K;A;T;A;V;A;G;L;S;R;P;A;G;P;A;E;L;S;R;L;G;E;P;L;I;V;S;E;D	A;Accession: I53715-8Qter	A;Accession: I67812/
F;P;2-512/Region: protein-tyrosine kinase lyn, splice form A #status predicted <MAT>	A;Accession: I53715-8Qter	A;Accession: I67812/
F;P;2-512/Binding site: protein-tyrosine kinase lyn, splice form A #status predicted <MAT>	A;Accession: I53715-8Qter	A;Accession: I67812/
F;P;2-512/Modifield site: myristylated amino end (Gly) (in mature form) #status predicted	A;Accession: I53715-8Qter	A;Accession: I67812/
F;P;3/3/Binding site: palmitate (Cys) (covalent) #status predicted	A;Accession: I53715-8Qter	A;Accession: I67812/
F;P;3/3/Active site: Lys #status predicted	A;Accession: I53715-8Qter	A;Accession: I67812/
F;P;3/97,508/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted	A;Accession: I53715-8Qter	A;Accession: I67812/
Query Match	Qy	6 SRRKSRLPSPLSSSVQGQPVTMEERSKATAVALGSFPGGGPAELSLRLGEPLTVSED 65
Best Local Similarity 27.4%	Db	38 SNKQRPVPVE SOLLPGQRFOAQDDEQGDIVVALXPYDGTHPDILSFKEGKMKVLEEH 96
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;	Qy	66 GDWWTLYLSEVGREYNIPSYHVGKV---SHGWLYEGLSREKAELLLPGPNGGAFLIR 121
Qy	Db	97 GEWWKAKSLSKREGIPSNVAKNTLEEWFFKDITRKDAEQLLAPNSAGAFLIR 156
Best Local Similarity 39.8%; Pred. No. 6.9e-23; Length 512;	Qy	97 GEWWKAKSLSKREGIPSNVAKNTLEEWFFKDITRKDAEQLLAPNSAGAFLIR 156
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;	Db	97 GEWWKAKSLSKREGIPSNVAKNTLEEWFFKDITRKDAEQLLAPNSAGAFLIR 156
Qy	Db	122 ESQTTRGSYSLSVRLSPASDRIRHYRTHCLDNGWLYISPRLTFSPLSQLVVDHYSELAD 181
6 SRRKSRLPSPLSSSVQGQPVTMEERSKATAVALGSFPGGGPAELSLRLGEPLTVSED 65	Qy	122 ESQTTRGSYSLSVRLSPASDRIRHYRTHCLDNGWLYISPRLTFSPLSQLVVDHYSELAD 181
38 SNKQRPVPVE SOLLPGQRFOAQDDEQGDIVVALXPYDGTHPDILSFKEGKMKVLEEH 96	Db	157 ESETLGFSLSVSDPQVDPMECDVTKYKRSLSNCGYYISPRTPCCISDMIKHYQKSD 216
66 GDWWTLYLSEVGREYNIPSYHVGKV---SHGWLYEGLSREKAELLLPGPNGGAFLIR 121	Qy	182 DICCLLKEPKCVLQRAGPLPK 202
97 GEWWKAKSLSKREGIPSNVAKNTLEEWFFKDITRKDAEQLLAPNSAGAFLIR 156	Db	217 GLCRRLJEKACT---SPKPQK 233
122 ESOTRGYSLSVRLSPASDRIRHYRTHCLDNGWLYISPRLTFSPLSQLVVDHYSELAD 181	Qy	RESULT 4
157 ESEPLKGFSLSVRLSPASDRIRHYRTHCLDNGWLYISPRLTFSPLSQLVVDHYSELAD 216	Db	A39719
182 DICCLLKEPKCVLQRAGPLPK 202	Qy	protein-tyrosine kinase (BC 2.7.1.112) lyn, long splice form - mouse
217 GLCRRLJEKACT---SPKPQK 233	Db	N;Contains: protein-tyrosine kinase lyn, short splice form
188 DCCCLLKEPKCVLQRAGPLPK 202	Qy	C;Species: Mus musculus (house mouse)
219 GLCRRLJEKACT---SPKPQK 233	Db	C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 03-Mar-2000
220 DCCCLLKEPKCVLQRAGPLPK 202	Qy	C;Accession: A39719; B39719; A39750; B39750

R;Stanley, E.; Ralph, S.; McBewen, S.; Boulet, I.; Holtzman, D.A.; Lock, P.; Dunn, A.R.  
 Mol. Cell. Biol. 11, 3399-3406, 1991  
 A;Title: Alternatively spliced murine lyn mRNAs encode distinct proteins.  
 A;Reference number: A39719; MUID:91260688; PMID:1710766  
 A;Accession: A39719  
 A;Molecule type: mRNA  
 A;Residues: 1-12 <STA1>  
 A;Cross-references: GB:M64608; NID:G198938; PID:AAA39470.1; PMID:9198939  
 A;Accession: B39719  
 A;Molecule type: mRNA  
 A;Residues: 1-24, 46-512 <STA2>  
 A;Cross-references: GB:M64608  
 A;Residues: 1-76 'F' ,78-160 'I' ,162-278 'L' ,280-390 'I' ,392-424 'D' ,426-512 <YII>  
 A;Cross-references: GB:M57696; NID:G198940; PID:AAA39471.1; PMID:9198941  
 A;Accession: C39750  
 A;Molecule type: mRNA  
 A;Residue number: A39750; MUID:91203957; PMID:2017160  
 A;Cross-references: GB:M67697; NID:G198942; PID:AAA39472.1; PMID:9198943  
 C;Superfamily: Protein-Tyrosine kinase src; protein kinase homology; SH3 hc  
 C;Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprote  
 F;1-12/Produced: protein-tyrosine kinase lyn, long splice form #status predicted <MAT>  
 F;1-14, 46-512/Produced: protein-tyrosine kinase lyn, short splice form #status predicted  
 F:129-226/Domain: SH3 homology <SH3>  
 F:249-504/Domain: SH2 homology <SH2>  
 F:255-261/Region: protein kinase ATP-binding motif  
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F:2/5/Active site: Lys #status predicted  
 F:399/508/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted  
 Query Match 27.2%; Score 367.5; DB 1; Length 512;  
 Best Local Similarity 40.3%; Pred. No. 1.2e-22;  
 Matches 91; Conservative 33; Mismatches 78; Indels 9; Gaps 3;  
 Qy 6 SRRRLDPSPQISSYQGGVTTMAERSATAVALGSPPAGGAELSLRGLPPLTIVED 65  
 Db 38 SNKQQRPVPEF-HLPGQFQTKDPEEQDIVVLYPDGHHDDLSFKKGKMKVLEEH 96  
 Qy 66 GDWWTVLSEVSGREYNIPSVHGVY---SHGMLEYGLSREKAELLILPQNGGAPLIR 121  
 Db 97 GEWWAKLSSKREKFPIFSNVYVARVNTLETFWDKFTRDQDERQLIAGNSAGAFLIR 156  
 Qy 122 BSOTRGGSYSLSVRLSRPASWDRHYRHCLDNGWLYISPRLTPEPSLQALVDHYSFLAD 181  
 Db 157 ESETURGSFSRSVRYDPMHGDKVHKYKTSLDNGYYTSPRITPEPCISDMKHYQKSD 216  
 Qy 182 DICCLKEPVQIQRAGPLCK 202  
 Db 217 GLCRLERAKCT---SPKPK 233  
 RESULT 5  
 TVHTIC  
 protein-tyrosine kinase (EC 2.7.1.112) hck - human  
 C;Species: Homo sapiens (man)  
 C;Date: 31-Dec-1989 #sequence revision 10-Nov-1995 #text\_change 11-Jun-1999  
 C;Accession: A27811; A27812; JCI149; C38268; S31103  
 R;Quintrell, N.; Lebo, R.; Varmus, H.; Bishop, J.M.; Pettenati, M.J.; Le Beau, M.M.; Diaz  
 Mol. Cell. Biol. 7, 2267-2275, 1987  
 A;Title: Identification of a human gene (HCK) that encodes a protein-tyrosine kinase and  
 A;Reference number: A27811; MUID:87257942; PMID:3496523  
 A;Accession: A27811  
 A;Molecule type: mRNA  
 A;Residues: 1-505 <QUI>  
 A;Cross-references: GB:M16591  
 A;Note: the codon given for 3-CYS (TCG) is inconsistent with the authors' translation  
 R;Ziegler, S.F.; Marth, J.D.; Lewis, D.B.; Perlmutter, R.M.

Mol. Cell. Biol. 7, 2276-2285, 1987  
 A;Title: Novel protein-tyrosine kinase gene (hck) preferentially expressed in cells of hematopoietic origin  
 A;Reference number: A27812; MUID:87257943; PMID:453117  
 A;Accession: A27812  
 A;Molecule type: mRNA  
 A;Residues: 1-505 <ZIE>  
 A;Cross-references: GB:M16592; NID:9183913; PID:AAA52644.1; PMID:9306833  
 R;Bradetzky, D.; Streibhardt, K.; Ruebsamen-Waigmann, H.  
 Gene 113, 275-280, 1992  
 A;Title: The genomic locus of the human hematopoietic cell specific protein tyrosine kinase  
 A;Reference number: JC1149; MUID:92241680; PMID:92241680  
 A;Accession: JJC149  
 A;Molecule type: DNA  
 A;Residues: 157-505 <HRA>  
 A;Cross-references: EMBL:X59741  
 A;Cross-references: T.P.; Alitalo, R.; Lehto, R.; Alitalo, K.  
 R;Partanen, J.; Maekelae, T.P.; Alitalo, K.; Lehto, R.; Alitalo, K.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990  
 A;Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.  
 A;Reference number: A38268; MUID:91062289; PMID:2247464  
 A;Accession: C38268  
 A;Status: nucleic acid sequence not shown; not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 362-417 <PAR>  
 C;Genetics:  
 C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phosphatase; protein kinase ATP-binding motif  
 A;Gene: GDN; HCK  
 A;Cross references: GDB:119303; OMIM:1423700  
 A;Intron Position: 20q11-20q12  
 A;Introns: 207/1; 258/1; 318/1; 343/3; 395/1; 439/1  
 C;Function:  
 A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 hc  
 C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phosphatase; protein kinase ATP-binding motif  
 F:2-505/Product: protein tyrosine kinase hck #status predicted <MAT>  
 F:764-112/Domain: SH3 homology <SH3>  
 F:123-220/Domain: SH2 homology <SH2>  
 F:239-497/Domain: protein kinase ATP-binding motif  
 F:247-255/Region: protein kinase ATP-binding motif  
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F:3/Binding site: palmitate (Cys) (covalent) #status predicted  
 F:269/Active site: Lys #status predicted  
 F:390/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted  
 Query Match 26.6%; Score 360.5; DB 1; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 4.6e-22;  
 Matches 7; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 Qy 12 PPSLLESSVQCGQPTMEARSATAVALGSPPAGGAELSLRGLPPLTIVEDDWTV 71  
 Db 40 PGPNHSNS---NTPGTRGAESEDILIVVVALDYEAETHHELSFQGDQMVYEECEWKA 96  
 Qy 72 LSEVGSGREYANPSPVGVK ---SHGWLYBGLSREKAELLILPQNGGAFALRBSQTAR 127  
 Db 97 RSLATEPKGIPSNVYARYDSLEFWFGISKDAEQLLAQNMLGSMFMDSETPK 156  
 Qy 128 GSYSLSYRLSPASWDRYRHYRHLQDNGWLYISPRLTPEPSLQALVDHYSFLAD 187  
 Db 157 GSYSLSYRLDPRQGDTVHVYKIRTLNDGGFYISPRSTSTLQELVDHYNKKGNGLCQKL 216  
 Qy 188 KEPCV 192  
 Db 217 SVPCM 221  
 RESULT 6  
 JC1321  
 protein-tyrosine kinase (EC 2.7.1.112) hck - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Sequence revision 10-Sep-1999 #text\_change 04-Feb-2000  
 C;Accession: JJC1321; S1874  
 R;Okano, Y.; Sugimoto, Y.; Fukukawa, M.; Matsui, A.; Nagata, K.; Nozawa, Y.  
 Biophys. Res. Commun. 181, 1137-1144, 1991  
 A;Title: Identification of rat cDNA encoding hck tyrosine kinase from megakaryocytes.  
 A;Reference number: JJJ1321; MUID:32109719; PMID:1764664

A;Accession: JO1321  
A;Molecule type: mRNA  
A;Residues: 1-503 <OKA>  
A;Cross-references: GB:74141; NID:G241436; PIDN:AAB20754.1; PID:9241437  
A;Experimental source: megakaryocyte  
R;Rema, V.; Sharup, G.  
submitted to the EMBL Data Library, December 1991  
A;Reference number: S18974  
A;Accession: S18974  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-50, 'V', 52-204, 'R', 206-305, 'T', 307-503 <REM>  
A;Cross-references: EMBL:X62345; NID:957581; PIDN:CAA44218.1; PID:957582  
C;Genetics:  
C;Keywords: hck  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH3 domain; kinase  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology; kinase-related transforming protein  
F:62:110/Domain: SH3 homology <SH3>  
F:121:218/Domain: SH2 homology <SH2>  
F:245:253/Region: protein kinase ATP-binding motif  
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F:3/Binding site: Palmitate (Cys) (covalent) #status predicted  
F:388/Binding site: phosphate (Tyz) (covalent) (by autophosphorylation) #status predicted  
Query Match Score 353.5; DB 1; Length 503;  
Best Local Similarity 37.1%; Pred. No. 1.7e-21;  
Matches 85; Conservative 36; Mismatches 77; Indels 31; Gaps 5;  
Qy 1 MGSLPSR--RKSLPSPLSSSVOCGPVME-----AERSKAT-A 37  
Db 1 MGCVKSRLFEGSKASKIENANQKGPVTPDPTSPKKLPGPNSTNSLPPFGESDTIV 60  
Qy 38 VALGSFPAGGAELSLRLGEPMTIVSDDGWWTVLSEVSGREYNIPTSYHGVV---SHG 93  
Db 61 VALYDEAIIHRDLFQKGQMWWBESGWKVKARSLATKREGTIPSYNTVARYNTSLTEE 120  
Qy 94 WLYEGHSREKAELLLPGNGGAPLIRESTOPRGYSYSLSVRLSRPASWDRIBHYRHL 153  
Db 121 WFFKGTSRKDAERHLLAQPGNLNGSMIRDETTKGYSYSLSVRDPOHGTPLPK 180  
Qy 154 DNGWLYISPRITFPSSLQALWDHYSLEADDICCLLKEPLQVLRQACPLPK 202  
Db 181 DSGGFYTISPSSTFESSQELVHYKKGDKQKLSVPCKV---SPKPQK 225

RESULT 7  
I37206 protein-tyrosine kinase (BC 2.7.1.112) blk - human  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-505 <RES>  
A;Cross-references: EMBL:Z333998; NID:g601951; PIDN:CAA83965.1; PID:9601952  
C;Genetics:  
A;Gene: BLK  
A;Cross-references: GDB:154111; OMIM:191305  
A;MAP position: 8p21-8p22  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH3 homology; kinase  
C;Keywords: ATP; blocked amino end; lipoprotein; myristylation; phosphotransferase; tyrosine kinase; kinase  
F:65:113/Domain: SH3 homology <SH3>  
F:124:220/Domain: SH2 homology <SH2>  
F:239:4-97/Region: protein kinase homology <KIN>  
F:247:255/Region: protein kinase ATP-binding motif

F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
E:269/Active site: Lys #status predicted  
Query Match Score 352.5; DB 2; Length 505;  
Best Local Similarity 43.6%; Pred. No. 2.1e-21;  
Matches 75; Conservative 24; Mismatches 68; Indels 5; Gaps 2;  
Qy 25 PVTMBAERSKATAVALGSFPAGGAELSLRLGEPMTIVSDDGWWTVLSEVSGREYNIPTSYHGVV---SHG 93  
Db 51 PPDEHEDKEDHFWVFLVYDYMAMNDLQWIKLQVKGKGTGDMWLLARLVTCPEGYPS 110  
Qy 85 VHVGKVIS ---HGMLYEGHSREKAELLLPGNGGAPLIRESTOPRGYSYSLSVRLSRPAA 140  
Db 111 NEVARVTELEMERNWFRSQRKEARQLAPINKAGSFPLIRESTOPRGYSYSLSVKT-DVTT 169  
Qy 141 SWDRIRHYRHCLDNGWLYISPRITFPSSLQALWDHYSLEADDICCLLKEPLQVLRQACPLPK 192  
Db 170 QGELIKHYKIRCLDEGGYYISPRIFPSLQALWDHYSLEADDICCLLKEPLQVLRQACPLPK 221

RESULT 8  
TYMSHC  
protein-tyrosine kinase (EC 2.7.1.112) hck - mouse  
A;Cross-references: GDB:154111; OMIM:191305  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1989 #sequence\_change 28-Jan-2000  
C;Accession: A27282; A39973  
R;Klemisz, M.J.; McRaeher, S.R.; Maki, R.A.  
Nucleic Acids Res. 15: 9600, 1987  
A;Title: Nucleotide sequence of the mouse hck gene  
A;Reference number: A27282; MUID:88067781; PMID:3684607  
A;Molecule type: mRNA  
A;Residues: 1-503 <KIN>  
A;Cross-references: GB:Y00487; NID:951209; PIDN:CAA68544.1; PID:951210  
R;Holtzman, D.A.; Cook, W.D.; Dunn, A.R.  
Proc. Natl. Acad. Sci. U.S.A. 84: 8325-8329, 1987  
A;Title: Isolation and sequence of a cDNA corresponding to a src-related gene expressed in  
A;Reference number: A39973; MUID:80687781; PMID:3317404  
A;Accession: A39973  
A;Status: Preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-503 <HOL>  
A;Cross-references: GB:J03023; NID:g192212; PIDN:AAA37305.1; PID:g309118  
C;Genetics:  
A;Gene: hck  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology; phosphotransferase; tyrosine kinase  
C;Keywords: ATP; autop phosphorylation; blocked amino end; lipoprotein; myristylation; phosphotransferase  
F:121:218/Domain: SH3 homology <SH3>  
F:237:495/Domain: protein kinase homology <SH2>  
F:245:253/Region: protein kinase ATP-binding motif  
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F:3/Binding site: pamitate (Cys) (covalent) #status predicted  
F:267:499/Active site: Lys #status predicted  
F:388:499/Binding site: phosphate (Tyr) (covalent) (by autop phosphorylation) #status predicted  
Query Match Score 352; DB 1; Length 503;  
Best Local Similarity 41.0%; Pred. No. 2.3e-21;  
Matches 82; Conservative 33; Mismatches 73; Indels 12; Gaps 4;  
Qy 11 LPSPLSLSVQQGPVME --AERSKAT-AVALGSFPAGGAELSLRLGEPMTIVSDDG 66  
Db 30 VPDPTSSKLGPNNNSMPPGFVEGSEDTIVVAVLYDEATHREDLSFKQCDOMVYLEEAG 89  
Qy 67 DWTVLVLBEVSREYNIPTSYHGVV ---SHGMLYEGHSREKAELLLPGNGGAPLIRESTOPRGYSYSLSVRLSRPAA 122  
Db 90 EWKAKSLATKEKGIPSNVARYNSLTELIEWFFGIGSRSRDAEHLAPCNMLGSEFMIRD 149  
Qy 123 SOTRGYSYSLSVRLSRPASWDRIRHYRHCLDNGWLYISPRITFPSSLQALWDHYSLEADDICCLLKEPLQVLRQACPLPK 182  
Db 150 SPTRGYSYSLSVRLSRPASWDRIRHYRHCLDNGWLYISPRITFPSSLQALWDHYSLEADDICCLLKEPLQVLRQACPLPK 209



Db 109 GPIPENFVAKANSLEPEPWFKNLRSKDAEQLLAPGNTGSFLTRESESTAGSFSLSYR 168  
 C;Species: Mus musculus (house mouse)  
 C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 03-Mar-2000  
 C;Accession: A23639; I57629; I74542;  
 R;Voronova, A.F.; Setton, B.M.  
 Nature 319, 682-685, 1986  
 A;Title: Expression of a new tyrosine kinase gene is stimulated by retrovirus promoter  
 A;Reference number: I48845; MUID:86146842; PMID:3081813  
 A;Accession: I48845  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-509 <VOR1>  
 A;Cross-references: GB:MI03533; PIDN:954813; PIDN:CAA27234..1; PID:G54814  
 B;March, J.D.; Peet, R.; Krebs, E.G.; Perlmutter, R.M.  
 Cell 143, 393-404, 1985  
 A;Title: A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpressed  
 A;Reference number: A23639; MUID:86079521; PMID:416464  
 A;Accession: A23639  
 A;Molecule type: mRNA  
 A;Residues: 1-232, 'VP', 285-509 <MAR>  
 A;Cross-references: GB:MI12056; PIDN:9198763  
 R;Voronova, A.F.; Adler, H.T.; Setton, B.M.  
 Mol. Cell. Biol. 7, 4407-4413, 1987  
 A;Title: Two lck transcripts containing different 5' untranslated regions are present in  
 A;Reference number: 157629; MUID:8814282; PMID:3501824  
 A;Accession: I57629  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-11 <VOR>  
 A;Cross-references: GB:MI18098; PIDN:9198766; PIDN:AAA39421..1; PID:g198767  
 R;Gavin, A.M.; Pawar, S.; March, J.D.; Perlmutter, R.M.  
 Mol. Cell. Biol. 8, 3058-3064, 1988  
 A;Title: Structure of the murine lck gene and its rearrangement in a murine lymphoma cell  
 A;Reference number: I57636; MUID:89096891; PMID:2850479  
 A;Accession: I7452  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-35, 'VR' <GAR>  
 A;Cross-references: GB:M22511; PIDN:9198768; PIDN:AAA39422..1; PID:g554186  
 C;Keywords: ATP; autophosphorylation; blocked amino end; kinase-related transforming protein  
 F;68-116/Domain: SH3 homology <SH3>  
 F;127-224/Domain: SH2 homology <SH2>  
 F;243-501/Domain: protein kinase homology <KIN>  
 F;251-259/Region: protein kinase ATP-binding motif  
 F;273/Active site: myristylated amino end (Gly) (in mature form) #status predicted  
 F;394-505/Binding site: phosphopeptide (Tyr) (covalent) (by autophosphorylation) #status predicted

## RESULT 11

I48845 Protein-tyrosine kinase (EC 2.7.1.112) lck, lymphocyte - mouse  
 N;Alternate names: p56; protein-tyrosine kinase lck  
 C;Species: Mus musculus (house mouse)  
 C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 03-Mar-2000  
 C;Accession: A23639; I57629; I74542;  
 R;Voronova, A.F.; Setton, B.M.  
 Nature 319, 682-685, 1986  
 A;Title: Expression of a new tyrosine kinase gene is stimulated by retrovirus promoter  
 A;Reference number: I48845; MUID:86146842; PMID:3081813  
 A;Accession: I48845  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-509 <VOR1>  
 A;Cross-references: GB:MI03533; PIDN:954813; PIDN:CAA27234..1; PID:G54814  
 B;March, J.D.; Peet, R.; Krebs, E.G.; Perlmutter, R.M.  
 Cell 143, 393-404, 1985  
 A;Title: A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpressed  
 A;Reference number: A23639; MUID:86079521; PMID:416464  
 A;Accession: A23639  
 A;Molecule type: mRNA  
 A;Residues: 1-232, 'VP', 285-509 <MAR>  
 A;Cross-references: GB:MI12056; PIDN:9198763  
 R;Voronova, A.F.; Adler, H.T.; Setton, B.M.  
 Mol. Cell. Biol. 7, 4407-4413, 1987  
 A;Title: Two lck transcripts containing different 5' untranslated regions are present in  
 A;Reference number: 157629; MUID:8814282; PMID:3501824  
 A;Accession: I57629  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-11 <VOR>  
 A;Cross-references: GB:MI18098; PIDN:9198766; PIDN:AAA39421..1; PID:g198767  
 R;Gavin, A.M.; Pawar, S.; March, J.D.; Perlmutter, R.M.  
 Mol. Cell. Biol. 8, 3058-3064, 1988  
 A;Title: Structure of the murine lck gene and its rearrangement in a murine lymphoma cell  
 A;Reference number: I57636; MUID:89096891; PMID:2850479  
 A;Accession: I7452  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-35, 'VR' <GAR>  
 A;Cross-references: GB:M22511; PIDN:9198768; PIDN:AAA39422..1; PID:g554186  
 C;Keywords: ATP; autophosphorylation; blocked amino end; kinase-related transforming protein  
 F;68-116/Domain: SH3 homology <SH3>  
 F;127-224/Domain: SH2 homology <SH2>  
 F;243-501/Domain: protein kinase homology <KIN>  
 F;251-259/Region: protein kinase ATP-binding motif  
 F;273/Active site: myristylated amino end (Gly) (in mature form) #status predicted  
 F;394-505/Binding site: phosphopeptide (Tyr) (covalent) (by autophosphorylation) #status predicted

QY 25.0% Score 338.5; DB 1; Length 509;  
 Best Local Similarity 39.1%; Pred. No. 3e-20;  
 Matches 77; Conservative 28; Mismatches 85; Indels 7; Gaps 2;

Db 119 LIRESQTRRGSSYSLSYVRLSPASWDRIRHYHCLDNGWLYISPLRTFPSSLQALVDHYSE 178  
 Db 152 LIRESQTRRGSSYSLSYVRLSPASWDRIRHYHCLDNGWLYISPLRTFPSSLQALVDHYSE 211

QY 179 LADDICCLIKEPCVCQR 195  
 Db 212 ASDGLCITKLSPCQCQTQR 228

RESULT 12

A39939 LIRESQTRRGSSYSLSYVRLSPASWDRIRHYHCLDNGWLYISPLRTFPSSLQALVDHYSE 178  
 protein-tyrosine kinase (EC 2.7.1.112) tk1 [similarity] - chicken  
 N;Alternate names: kinase-related transforming protein (tk1); T-cell surface antigen ass:  
 C;Species: Gallus gallus (chicken)  
 C;Accession: A42126; A39939  
 C;Date: 16-Jun-2000 #sequence\_revision 16-Jun-2000 #text\_change 16-Jun-2000  
 R;Chow, L.M.; Ratcliffe, M.J.; Ruebsamen-Waigmann, H.  
 Mol. Cell. Biol. 12, 1226-1233, 1992  
 A;Title: tk1 is the avian homolog of the mammalian lck tyrosine protein kinase gene.  
 A;Reference number: A42126; MUID:92186654; PMID:1545804  
 A;Accession: A42126  
 A;Molecule type: mRNA  
 A;Residues: 1-88 <CHO>  
 A;Cross-references: GB:MB5043  
 A;Experimental source: thymus, spleen  
 A;Note: sequence extracted from NCBI backbone (NCBIN:88831, NCBIPI:88833.)  
 R;Treblettard, K.; Mullins, J.; Bruck, C.; Ruebsamen-Waigmann, H.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 8778-8782, 1987  
 A;Title: Additional member of the protein-tyrosine kinase family: the src-and lck-related  
 A;Accession: A39939  
 A;Molecule type: mRNA  
 A;Residues: 52-507 <STR>  
 A;Cross-references: GB:J03579; PIDN:9A49081..1; PID:9212713  
 C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phosphorylation  
 F;65-114/Domain: SH3 homology <SH3>  
 F;125-222/Domain: SH2 homology <SH2>  
 F;241-499/Region: protein kinase homology <KIN>  
 F;249-257/Region: protein kinase ATP-binding motif  
 F;392,503/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted:  
 Query Match 24.6%; Score 333; DB 1; Length 507;  
 Best Local Similarity 43.2%; Pred. No. 8.7e-20;  
 Matches 70; Conservative 27; Mismatches 61; Indels 4; Gaps 1;

QY 38 VALGSFPAAGPAAELSLURGEPLTTIVSGLWTVLSSBVGSGREYNPSVHYRKVS---HG 93  
 Db 65 VAIYDVEPTHDGDGLGMQEKEVRLVESEGWTAQSITLGQEGLIPNPFYAMVNSLEPEP 124

QY 94 WIYEGLSREKABELLPLPGNPAGAFLIRETSQTRGSYSLSVLSPASWDRIRHYHCL 153  
 Db 125 WFPANLSKRNAEARLLSGNTHSFLRESETSGKSYSLSVLSPASWDRIRHYHCL 184

QY 154 DNGWLYISPLRTFPSSLQALVDHYSE 178; Pred. No. 1e-20;  
 Db 185 DNGYYIISPRVTFSSLHLVETYSSSSDGLCITRLGRKCRCRQK 226

RESULT 13

B49114 DNGWLYISPLRTFPSSLQALVDHYSE 178; Pred. No. 1e-20;  
 protein-tyrosine kinase (EC 2.7.1.112) fyk - Pacific electric ray  
 C;Species: Torpedo californica (Pacific electric ray)  
 C;Accession: B49114  
 R;Scope, S.L.; Hogenauer, R.L.  
 J. Biol. Chem. 268, 25152-25161, 1993  
 A;Title: Molecular cloning of two abundant protein tyrosine kinases in Torpedo electric  
 A;Reference number: A49114; MUID:94043386; PMID:8227079  
 A;Accession: B49114  
 A;Status: preliminary  
 A;Molecule type: mRNA

A;Residues: 1-539 <SWO>  
 A;Cross-references: GB:U01350  
 C;Superfamily: Protein-tyrosine kinase src; protein kinase homology; SH3 h  
 C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; Phc  
 F;91-140/Domain: SH3 homology <SH2>  
 F;151-248/Domain: SH2 homology <SH2>  
 F;279-287/Region: protein kinase ATP-binding motif  
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F;422,533/Binding site: Lys #status predicted  
 F;422,533/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 24.0% Score 325; DB 2; Length 539;  
 Best Local Similarity 41.0% Pred. No. 4.3e-19;  
 Matches 75; Conservative 27; Mismatches 67; Indels 14; Gaps 4;

Qy 15 SISSSVQG-QPYNTMEADRSKATAVALGSFAAGPPELSRLGEPLTV-SEDDWWTVL 72  
 Db 74 SLTRSGGGTGTGTM-----FIALNNYDARTEDDLTRKGFERHIIINSSEGDWNEAR 125  
 Qy 73 SEVSGREINNPSVHVGKV---SHGWLYEGGSREKABELLLPGNPGGAELIRESQTERG 128  
 Db 126 SLITGSTCYPSNVYAPVDSIQAEENYFGKNGRKDKTERMILCPGNPRGTFLIRESETKG 185  
 Qy 129 SYSLSVRUPRASWDREHYRHFLCDLNGWLYISPRLTFSQQALNDHYSBELADDICLRL 188  
 Db 186 AYSLSLIRDWDVEKGHDVHYKIRKLDDNGYYITTRAQPERVPOLVHYTERAAGLCRLV 245  
 Qy 189 EPC 191  
 Db 246 VPC 248

RESULT 14  
 TWF99  
 protein-tyrosine kinase (EC 2.7.1.112) Yes - avian sarcoma virus Y73  
 C;Species: avian sarcoma virus Y73  
 A;Note: host Gallus gallus (chicken)  
 C;Date: 27-Nov-1995 #sequence\_revision 27-Nov-1995 #text\_change 23-Feb-1997  
 C;Accession: A00633  
 R;Kitamura, N.; Kitamura, A.; Toyoshima, K.; Hirayama, Y.; Yoshida, M.  
 Nature 297, 205-208, 1982  
 A;Title: Avian sarcoma virus Y73 genome sequence and structural similarity of its transi  
 A;Reference number: A00633; MUID:82195528; PMID:6281655  
 A;Molecule type: genomic RNA  
 A;Residues: 1-528 <KIT>  
 C;Comment: This protein is synthesized as a gag-ges polyprotein.  
 C;Genetics:  
 A;Gene: Yes  
 C;Superfamily: Protein-tyrosine kinase src; protein kinase homology; SH3 h  
 C;Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; trans  
 F;88-137/Domain: SH3 homology <SH3>  
 F;255-523/Region: protein kinase homology <SH2>  
 F;292/Active site: Lys #status predicted  
 F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 23.7% Score 320; DB 1; Length 528;  
 Best Local Similarity 31.5% Pred. No. 1.1e-18;  
 Matches 91; Conservative 43; Mismatches 105; Indels 50; Gaps 9;

Qy 2 GSPLSPRRKSPLSPSPLSSVYQGQPVTPMAERSKATAVALGSFPAAGPAAELSRLGEPEUTI 61  
 Db 62 GGASSSFSAVSPS-YPSLTGGTV-----FVALVYEARTDDISPKKGERQI 110  
 Qy 62 VSE-DGDNWNTVLSEVSGREINYIPSHYGKV---SHGMWLYEGLSREKABELLILPGNPGG 116  
 Db 111 INNTEGDWWEARSIAATGKIGYIFSNYTAPOSIQAEBWFRGRMKDAEPLLNPGNORG 170  
 Qy 117 AFLIREQSTRGSYSLSVLRSRASNDR----HYRHLDDNGWLYISPRLTFSPLQA 171

Db 171 IFLYRESETKGAVSLSIR----DWDEVRGDNVRYKIRKLDDNGYYITTRAQFESLSQK 225  
 Qy 172 LDVHYSLEADDICLRLKEPC-----VLRDRAGPILPGKNUPLPTVQR----- 212  
 Db 226 LVKTHREHAGLCKLTIVCPTVKPQTQGLADWEPRESSURLEVKGQCFGEVWMGTR 285  
 Qy 213 ---TPLNKXELDSSLTSSEAATGEESILSEGUREST-SFTISLANDBAV 256  
 Db 286 WNGTTRKVAATKLXLTGMMPEAFQEAQMKKURHDKLKVPLVAVVSBPI 334

RESULT 15  
 A45501  
 protein-tyrosine kinase (EC 2.7.1.112) Yes [similarity] - African clawed frog  
 N;Alternate names: kinase-related transforming protein (yes)  
 C;Species: Xenopus laevis (African clawed frog)  
 C;Date: 16-Jun-2000 #sequence\_revision 16-Jun-2000 #text\_change 16-Jun-2000  
 C;Accession: A45501; S08517  
 R;Steele, R.E.; Irwin, M.Y.; Krudsen, C.L.; Collett, J.W.; Fero, J.B.  
 Oncogene Res. 1, 223-233, 1989  
 A;Title: The yes proto-oncogene is present in amphibians and contributes to the maternal  
 A;Reference number: A45501  
 A;Accession: A45501  
 A;Molecule type: mRNA  
 A;Residues: 1-537 <STE>  
 A;Cross-references: GB:X14377  
 R;Steele, R.E.; Irwin, M.Y.; Knudsen, C.L.; Collett, J.W.; Fero, J.B.  
 submitted to the EMBL Data Library, February 1989  
 A;Accession number: S08517  
 A;Accession: S08517  
 A;Molecule type: mRNA  
 A;Residues: 1-250, S', 252-537 <ST2>  
 A;Cross-references: EMBL:X14377; PID:965272; PIDN:CAA32551-1; PID:965273  
 C;Genetics:  
 A;Gene: Yes  
 C;Superfamily: Protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 hc  
 C;Keywords: ATP; autophosphorylation; blocked amino end; kinase-related transforming prot  
 P;92-141/Domain: SH3 homology <SH3>  
 P;152-249/Domain: SH2 Homology <SH>  
 P;269-527/Domain: Protein kinase homology <KIN>  
 P;277-285/Region: protein kinase ATP-binding motif  
 P;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 P;99/Active site: Lys #status predicted  
 P;420,531/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted  
 Query Match 23.6% Score 319.5; DB 1; Length 537;  
 Best Local Similarity 33.2%; Mismatches 38; Indels 39; Gaps 7;  
 Matches 84; Conservative 38; Mismatches 92; Indels 39; Gaps 7;  
 Qy 38 VALGSPFAGGPAPLSLRLGGPLTIVSE-DGDMWTVLSEVSGREINYIPSHYGKV---SH 92  
 Db 91 VALDYEARTEEDSFGRGERQFQINNTGDMWEARSIAATGKIGYIFSNYTAPODIQE 150  
 Qy 93 GMVLEGLSREKABELLILPGNPGGAEILIRESGOTRGSYSLSVRLSPASWDRIR----H 147  
 Db 151 EWRGRMGRKDARLLINGNPGNRGTFLVRESSTTKGAYSLSIR----DIDEVRGDNYRHK 205  
 Qy 148 YRIHLDNGWLYISPRLTFSPLQLDVHYSSELADDICLRLKEPC-----VLRQACP 198  
 Db 206 YKIRKLDDNGYIITRAQFESLQKLVYHYSSELADDICLRLKEPC-----VLRQACP 265  
 Qy 199 LPGRDIPLEVTRQ-----TPIANWKEDSSLLFSEATGEBSLLSEGLRS 244  
 Db 266 IPRESLRLVYKLGCGFCGEWIGWTNGNTTKVAKTULKPGTMMPAFQEAQMCKLDRHK 325  
 Qy 245 L-SFYISLANDBAV 256  
 Db 326 LVPLYAVVSEEP1 338

Search completed: February 20, 2004, 09:33:04  
 Job time : 46 secs



RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., RA Hucke E., Hunt A.R., Hunt S.E., Johnson C.M., Johnson D., Johnson D., RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., RA Levensaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., RA Marsh V.L., Martin S.L., McConaughie L.J., McLay K., McMurray A.A., RA Milne S.A., Mistri D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillemore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Sims S., RA Rice C.M., Ross M.T., Scott C.B., Shownkeen R., Shuh H.K., Sodergren C., RA Skuce C., Smith M.L., Soderlund C., Sulston J.E., Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., RA Tracey A., Tronman A.C., Vaidin M., Wall J.M., Wallis J.M., Williams L., Williams S.A., RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;  
 RA "The DNA sequence and comparative analysis of human chromosome 20.",  
 RL Nature 414:665-671 (2001).  
 RN [5].

RP SEQUENCE FROM N.A.  
 RC TISSUE-PROstate;  
 RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Schuler G.D.,  
 RA Autchouk S.F., Zeaberg B., Bustow R., Blat N.K., Shemesh C.M., Shemesh C.M.,  
 RA Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soeda M.F., Casavant T.L., Schatz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Prange C.,  
 RA Raha S.S., Loqueland N.A., Peters C., Abramson R.D., Mullihay S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E., Lu J., Gibbs R.A.,  
 RA Fahey J., Helton E., Ketteman M., Madan A.J., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schein J.B., Schein J.B., Jones S.J.M., Marr M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [6].

RP CHARACTORIZATION;  
 RX MEDLINE=22013997; PubMed=11891219;  
 RA Pandey A., Ibarrola N., Kratchmarova I., Fernandez M.M., Fernandez M.M.,  
 RA Constantinescu S.N., Ohara O., Savasdossos S., Lodish H.F., Mann M., Mann M.,  
 RT "A novel Src homology 2 domain-containing molecule, Src-like adapter protein-2 (SLAP-2), which negatively regulates T cell receptor signaling";  
 RT "FUNCTION: Adapter protein, which negatively regulates T-cell receptor induced activation of nuclear factor of activated T-cells. May act by linking signaling proteins such as ZAP70 and CD32 via CBL dependent degradation of signaling proteins."  
 CC "SUBUNIT: Interacts with phosphorylated proteins ZAP70 and CD32 via its SH2 domain (by similarity). Interacts with phospho-activated CBL via its C-terminal domain."  
 CC "SUBCELLULAR LOCATION: Cytoplasmic. Isoform 1 is localized to the plasma membrane and intracellular vesicles, including late endosomal vesicles. Isoform 2 may be cytoplasmic and is not localized to membranes."  
 CC "ALTERNATIVE PROTEIN PRODUCTS:  
 Event=Alternative splicing; Named isoforms=2;  
 IsoID:Q9HQQ3-1; Sequence=Displayed;  
 Note=Isoform 3 is produced by alternative initiation at Met-27 of isoform 1;  
 Name=2; Synonyms=p28;  
 IsoID:Q9HQQ3-2; SLAP-2-v; MANS-v;  
 Note=Isoform 4 is produced by alternative initiation at Met-27 of isoform 2;  
 Event=Alternative initiation;

CC "Event=Alternative splicing; Named isoforms=2;  
 IsoID:Q9HQQ3-1; Sequence=Displayed;  
 Note=Isoform 3 is produced by alternative initiation at Met-27 of isoform 1;  
 Name=2; Synonyms=p23;  
 IsoID:Q9HQQ3-2; SLAP-2-v; MANS-v;  
 Note=Isoform 4 is produced by alternative initiation at Met-27 of isoform 2;  
 Event=Alternative initiation;

CC Comment=4 isoforms, 1/p28 (shown here), 2/p23, 3/p25 and 4/p20, are produced by alternative initiation at Met-0 and Met-27;  
 CC TISSUE SPECIFICITY: Predominantly expressed in immune system, with highest levels in peripheral blood leukocytes. Expressed in spleen, thymus and lymph nodes. Expressed in T-cells as well as monocytes, and at low level in B-cells. Also detected in placenta, prostate, skin, retina and colon.  
 CC DOMAIN: The loss of the C-terminal domain partially abolishes the inhibitory function, but can be partially compensated by higher level of protein expression.  
 CC PTM: May be phosphorylated (By similarity).  
 CC SIMILARITY: Contains 1 SH2 domain.  
 CC SIMILARITY: Contains 1 SH3 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).  
 CC

CC DR AF26353; PAB2904\_1;  
 CC DR AF290985; PAB38197\_1;  
 CC DR AF290986; PAB38198\_1;  
 CC DR AF025645; BAB15201\_1;  
 CC DR AF031622; CAB4445\_1; ALT\_TERM.  
 CC DR AL050318; CAB7545\_1;  
 CC DR BC042041; AAH42041\_1;  
 CC DR BC042041; AAH42041\_1;  
 CC DR BC06239; BLACK;  
 CC DR Genew; HGNC:17329; SLA2.  
 CC DR InterPro; IPR000980; SH2.  
 CC DR InterPro; IPR001052; SH3.  
 CC DR InterPro; IPR001052; SH3.  
 CC DR PF00017; SH2; 1.  
 CC DR PF00018; SH3; 1.  
 CC DR PR00401; SH2DOMAIN.  
 CC DR PD00093; SH2; 1.  
 CC DR SMART; SM00252; SH2; 1.  
 CC DR SMART; SM00326; SH3; 1.  
 CC DR PROSITE; PS50001; SH2; 1.  
 CC DR PROSITE; PS50002; SH3; 1.  
 CC KW Membrane; SH2 domain; SH3 domain; Myristylation; Phosphorylation;  
 KW Alternative initiation; Alternative splicing; Lipoprotein.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT CHAIN 1 260 SRC-LIKE-ADAPTER 2, ISOFORM 3 AND ISOFORM 4.  
 FT CHAIN 27 260 SRC-LIKE-ADAPTER 2, ISOFORM 3 AND ISOFORM 4.  
 FT INT\_MET 27 27 FOR ISOFORM 3 AND ISOFORM 4.  
 FT DOMAIN 31 91 SH2.  
 FT DOMAIN 93 190 SH2\_C-TERMINAL.  
 FT DOMAIN 194 260 N-myristoyl glycine (By similarity).  
 FT LIPID 1 1 LADDICILLKEPCVILQRAGPUPGKDPLPPTV->GWAP  
 FT VARSPLIC 178 209 WQGYTPPCDCAEDTQLERAGQLPPVF (in isoform 2).  
 FT VARSPLIC 210 260 Missing (In isoform 2).  
 FT MUTAGEN 1 1 /FTId=vsp\_007240.  
 FT SEQUENCE 260 AA; 28454 MW; A402C03449261B3D CRC64;  
 FT SEQUENCE 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC Query Match 99.2%; Score 1342; DB 1; Length 260;  
 CC Best Local Similarity 99.6%; Pred. No. 3.7e-105;

CC Qry 2 GSLSRSRRLPSLSSSVQGPQPVMEAEERSKATAVALGSPPAGPALSIRLGEPLTI 61  
 CC 1 GSLSRSRRLPSLSSSVQGPQPVMEAEERSKATAVALGSPPAGPALSIRLGEPLTI 60  
 CC Note=Isoform 4 is produced by alternative initiation at Met-27 of isoform 2;  
 CC Note=Isoform 3 is produced by alternative initiation at Met-27 of isoform 1;

CC 62 VSEGDWIVTVLSEVSREYNNIFSVHGVSKWGLSREKAELLPGNPGCAFIR 121

Db 61 VSEDDKWTVLSEVSREYNTIPSVIYAKVSHGWLTTEGLSSKEAELLLPGNPGAFIR 120  
 Db 122 ESOTRGSYSLSVRLSPASDRIHYRFLCDNGWLYISPRITPSQLQVLDYSELAD 181  
 Qy 121 ESOTRGSYSLSVRLSPASDRIHYRFLCDNGWLYISPRITPSQLQVLDYSELAD 180  
 Db 182 DCCCLLKEKPCVYQAGPLPGKIDPLPVTVOPTPLWKEDLSSLEATGEBSLSEL 241  
 Qy 181 DCCCLLKEKPCVYQAGPLPGKIDPLPVTVOPTPLWKEDLSSLEATGEBSLSEL 240  
 Db 242 RESSLFYTISLNDEAVSLDDA 261  
 Db 241 RESSLFYTISLNDEAVSLDDA 260

RESULT 2

ID SLA2\_MOUSE STANDARD; FRT; 258 AA.  
 AC QBR410; QBC0K2; Q8V1Z9; Q9D1Z9; 42. Created)  
 DT 10-OCT-2003 (Rel. 42. Last sequence update)  
 DE SRC-like-adapter 2 (src-like adapter protein-2) (SLAP-2).  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN 111  
 SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), ALTERNATIVE INITIATION,  
 RP CHARACTERIZATION, FUNCTION, MYristoylation, PHOSPHORYLATION,  
 RP INTERACTION WITH ZAP70 AND CBL, AND MUTAGENESIS OF GLY-1; MET-26 AND  
 RP ARG-119.  
 RX MEDLINE=22022000; PubMed=12024406;  
 RA Loreto M.P., Berry D.M., McGlade C.J.,  
 RT "Functional cooperation between c-Cbl and Src-like adaptor protein 2  
 in the negative regulation of T-cell receptor signaling.";  
 RL Mol. Cell. Biol. 22:441-4525 (2002).

[2] SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, MYristoylation  
 RP INTERACTION WITH CBL; ZAP70 AND CD32, AND MUTAGENESIS OF GLY-1; PRO-81  
 RP AND ARG-119.  
 RX MEDLINE=2201997; PubMed=11891219;  
 RA Pandey A., Ibarrola N., Kratchmarova I., Fernandez M.M.,  
 RA Constantinescu S.N., Ohara O., Sawasaki-Koiso S., Loelish H.F., Mann M.;  
 RT "A novel SRC homology 2 domain containing molecule Src-like adapter  
 protein-2 (SLAP-2), which negatively regulates T cell receptor  
 signaling.";  
 RL J. Biol. Chem. 277:19131-19138 (2002).

[3] SEQUENCE FROM N.A. (ISOFORM 1)  
 RP SPAIN=057BL/6J, and NOD; TISSUE=Retina, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nakaio I., Osato N., Saito R., Suzuki H., Yamakawa I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bulk C., Hume D.A., Quackenbush J.,  
 RA Scirimil L.M., Kanpin A., Matsuda H., Batyalov S., Beisele K.W.,  
 RA Blaire J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dregani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gatsboidi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Keddiessi R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenthard B., Lyons P.A.,  
 RA Maggio D.R., Mallais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Parvan W.J., Peretz G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed J.J., Reid J., Ring J., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Wattanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Kojino H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa T., Itoh M., Kagawa I.,  
 RA Hashizume W., Imori K., Ishii Y., Itoh T., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Barrey E., Hayashizaki Y.  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 CC -!- FUNCTION: Adapter protein, which negatively regulates T-cell  
 CC receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced  
 CC activation of nuclear factor of activated T-cells. May act by  
 CC linking signaling proteins such as ZAP70 with CBL, leading to a  
 CC CBL dependent degradation of signaling proteins.  
 CC -!- SUBUNIT: Interacts with phosphorylated CBL via its C-terminal  
 CC domain. Interacts with phosphorylated Proteins ZAP70 and CD32 via  
 CC its SH2 domain.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; localized to the plasma  
 CC membrane and intracellular vesicles, including late endosomal  
 CC vesicles.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative initiation;  
 CC Comment=2 isoforms, 1/p28 (shown here) and 2/p25, are produced  
 CC by alternative initiation at Met-0 and Met-26;  
 CC -!- TISSUE SPECIFICITY: Mainly expressed in immune system. Highly  
 CC expressed in spleen and thymus and expressed at intermediate  
 CC levels in lung. Not expressed in liver, heart and brain. Isoform 1  
 CC is predominant in lung and spleen, while isoform 2 is predominant  
 CC in thymus.  
 CC -!- DOMAIN: The loss of the C-terminal domain partially abolishes the  
 CC inhibitory function.  
 CC -!- PTM: May be phosphorylated.  
 CC -!- MISCELLANEOUS Ref-1 confirmed the alternative initiation by  
 CC mutating the Met in position -1 to Val, and showed that isoform 1  
 CC is abolished in favor of isoform 2.  
 CC -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC DR EMBL; AF087457; AAB38196; 1.  
 CC DR EMBL; AF034990; AAB16403; 1.  
 CC DR EMBL; AK020837; BAB32223; 1.  
 CC DR EMBL; AK030877; BAC22718; 1.  
 CC DR EMBL; AK088672; BAC40495; 1.  
 CC DR HSSP; P06239; L1LK.  
 DR MGII; MGII-1925059; S1a2.  
 DR GO; GO-0016023; C:cytoplasmic vesicle; IDA.  
 DR GO; GO-000570; C:late endosome; IDA.  
 DR GO; GO-0005886; C:plasma membrane; IDA.  
 DR GO; GO-0005515; F:protein banding; IDA.  
 DR GO; GO-0042110; P:T-cell activation; IDA.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH2.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR Prodrom; PD000093; SH2; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00326; SH2; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR KW Immune response; Membrane; SH2 domain; SH3 domain; Myristate;  
 KW Phosphorylation; Alternative initiation; Lipoprotein.  
 ET INIT-MET 0 0  
 PT CHAIN 1 258 SRC-LIKE-ADAPTER 2 ISOFORM 1.

Query Match 75.6%; Score 1023; DB 1; Length 258;  
 Best Local Similarity 79.3%; Pred. No. 1. 8e-7;  
 Matches 207; Conservative 16; Mismatches 34; Indels 4; Gaps 3;

Qy	2	GSLPSRRSLSPSBLSSVQGQPVTMEAEERKATAVALGEGPAGCBAELSLRLGEPPLT	61
Db	1	GSLRSRGTL-SSPPPSGQDQPSPVSPQPERHKVTAVALGSAGQARLSLRLGEPPLT	59
Qy	62	VSEGDGWTVLSEYSGREYNIPSYHV/GKVS/HWLYEGSLSREKAELLPPNPNGGAFLIR	121
Db	60	ISEGDWTVQVSEYSGREYHMPSVYVAKHNLVEGSLREKAELLPPNPNGGAFLIR	119
Qy	122	ESCPQRGSYSLSLRSRSPASWDIHRVHRIHDNGMLYISPRLEPSLQALVDHYSELAD	181
Db	120	ESCPQRGCGSYSLSLRSRSPASWDIHRVHRIQDNGMLYISPRLEPHVHYSBLAD	179
Qy	182	DICCLLKEPQLVQPAQGPQDPLPVTVQRPLNWLDSLLFSAA-TGEESELSEG	240
Db	180	GICCPLRPEVQLQKLGPKDTPPPVTVPTSSLNWKLDLRSLLFEPASSEASLISEG	239
Qy	241	LRESLSFYISLNDEAVSLDDA	261
Db	240	LRESLSYYSLAED--PLDDA	258

RESULT 3  
 SLA-RAT STANDARD; PRT; 312 AA.

SLA1-RAT  
 AC P59722;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE SRC-Like-adaptor (Src-like-adaptor protein 1).  
 GN Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1] SEQUENCE FROM N.A.  
 RP Alifragis P., Molnar Z., Parravolas J.;  
 RA Submitted (JAN 2003) to the EMBL/GenBank/dbJ databases.

-!- FUNCTION: Adapter protein, which negatively regulates T-cell receptor (TCR) signaling. Inhibits T cell antigen-receptor induced activation of nuclear factor of activated T-cells. Involved in the negative regulation of positive selection and mitosis of T-cells. May act by linking signaling proteins such as Zap70 with CBL, leading to a CBL dependent degradation of signaling proteins (By similarity).  
 -!- SUBUNIT: Homodimer. Interacts with phosphorylated CBL, SYK and LAT. Homodimerization and interaction with phosphorylated CBL occurs via its C-terminal domain. Interacts with phosphorylated PDGFRA and EPHA2. Interacts with proteins ZAP70; CD32, YAV1 and LCP2 via its SH2 domain (By similarity).  
 -!- SUBCELLULAR LOCATION: Cytoplasmic; localizes with endosomes (By

similarity). C-terminal domain is essential for the homodimerization and the interaction with CBL. While the interaction with CBL is apparently mediated via the hydrophobic region of this domain, the highly charged region is apparently required for the homodimerization (By similarity).  
 CC PTM: Phosphorylated (By similarity).  
 CC SIMILARITY: Contains 1 SH2 domain.  
 CC SIMILARITY: Contains 1 SH3 domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).  
 CC DR EMBL; AY217759; AAC061134.1; --.  
 CC DR InterPro; IPR000090; SH2.  
 CC DR InterPro; IPR0001452; SH3.  
 CC DR Pfam; PF00017; SH2; 1.  
 CC DR Pfam; PF00018; SH3; 1.  
 CC DR PRINTS; PR00048; SH2DOMAIN.  
 CC DR PRODOM; PD000093; SH2; 1.  
 CC DR SMART; SM00251; SH2; 1.  
 CC DR SMART; SM00326; SH3; 1.  
 CC DR PROSITE; PS50001; SH2; 1.  
 CC DR PROSITE; PS50001; SH3; 1.  
 CC DR SH2 domain; SH3 domain; Phosphorylation.  
 CC DR DOMAIN 38 98 SH3.  
 CC DR DOMAIN 100 191 SH2.  
 CC DR DOMAIN 206 312 AA; SLA C-TERMINAL.  
 CC SEQ 312 AA; A05412D39AB68388 CRC64;  
 CC Query Match 37.0%; Score 501; DB 1; Length 312;  
 CC Best Local Similarity 41.3%; Pred. No. 1. 1e-34;  
 CC Matches 112; Conservative 41; Mismatches 90; Indels 28; Gaps 6;  
 CC Qy 11 LPSPPLSSVQGQG----PVTIMEAERSKATAV----ALGSFPAGGPAELSRLQGEP 58  
 CC Db 5 LPGPSTSRGEKEMNSMKSTPAPERPLSNTEGLESDFLAVLNDYPSPDISPPIFRREEK 64  
 CC Qy 59 LTVSEDGDWRWTVLSEVSGREYNIPSPVHVGVSHGLNEGLSREKAELLPPNPNGGAF 118  
 CC Db 65 LRVISDEGGWIKVAKSLSTGRSYESIYPGICVARYHGWLFLGLRDKAEEELQLPDKTSF 124  
 CC Qy 119 LIRESTORRGSYSVLSRSPASDRYRIRHICLNDNGMLYISPRLEPSLQALVDHSE 178  
 CC Db 125 MIRESETKGPSVLSVR----HRQVKGYRIFPLPNATYYISPRLETCLEDVTHSE 178  
 CC Qy 179 LADDICCLIKEKPCPVLR----AGPLPGKDIPLPVTVQRTPLNKELDSLLEFAATG- 232  
 CC Db 179 VADGLCQLVTPCLQNTPAFTAQSPTSPGSPVTLRQTKFDNKRVSSLQEFGEGANP 238  
 CC Qy 233 ---EBSLISEGLRESLSEFYSL-NDEAVSLD 259  
 CC Db 239 LRDVDESFSYGLRESLASYSLTGDSNSFD 269

RESULT 4  
 SIAL MOUSE STANDARD; PRT; 280 AA.  
 CC ID SIAL MOUSE; STANDARD; PRT; 280 AA.  
 CC AC Q60895; Q8C9QB; Q8CBEP; Q9QZX8;  
 CC DT 10-OCT-2003 (Rel. 42, Created)  
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 CC DE SRC-1-like-adaptor (Src-like-adaptor Protein 1) (mSLAP).  
 CC GN SLA OR SLAP OR SLAPI.  
 CC OS Mus musculus (Mouse).  
 CC OC Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus.  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 CC OK

- [1] SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND INTERACTION WITH EphA4.  
RP Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RP Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RP Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RC Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,  
RX MEDLINE=99370243; PubMed=7543898;  
RA Pandey A., Duan H., Dixit V.M.;  
RT "Characterization of a novel Src-like adapter protein that associates with the Eck receptor tyrosine kinase.";  
RT J. Biol. Chem. 270:19201-19204 (1995).  
RL [2]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RP MEDLINE=20093985; PubMed=10630289;  
RA Carrier A., Nguyen C., Victorero G., Granjeaud S., Rocha D.,  
RA Malissen B., Miazek A., Ferrier P., Malissen M., Naquet P.,  
RT "Differential gene expression in CD3epsilon- and RAG1-deficient thymocytes: definition of a set of genes potentially involved in thymocyte maturation.";  
RT Immunogenetics 50:255-270 (1999).  
RL [3]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=LPS, and ISS. (ISOFORMS 1 AND 2).  
RX MEDLINE=1363810; PubMed=11471062;  
RA Shringher M.A., Thompson J., Conroy C., Xu Y., Yang F., Canniff J.,  
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;  
RT "High-throughput sequence identification of gene coding variants within alcohol-related QTLs";  
RT Mamm. Genome 12:657-663 (2001).  
RL [4]  
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RP STRAIN=C57BL/6J; TISSUE=cerebral cortex, and Thymus;  
RX MEDLINE=23354683; PubMed=12456851;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nakanishi R., Saito N., Saito R., Suzuki H., Yamakata I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hassegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapana R., Matuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Frazer K.S.,  
RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Godzik A.,  
RA Gaasterland T., Gustafsson M., Girosi C., Gottschall P., Gough J.,  
RA Grimmond S., Gustincich S., Hirakawa N., Jackson J.J., Jarvis E.D.,  
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kuroki I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglione D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G.,  
RA Petrusova N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultan R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Varardo R., Wagner L., Wahlestedt C., Wang Y., Watansabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carrinci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imoto K., Ishii Y., Itoh M., Kagewa I.,  
RA Miyazaki A., Sakai K., Sesaki D., Shibata K., Shinagawa A.,  
RA Yasunishi E., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=mammary gland;  
RX MEDLINE=2288257; PubMed=12477332;  
RA Straubhaar R.L., Pengold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shevchenko A.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buter K.H., Schaeffer C.F., Bhat N.K.,  
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh F.P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Loqueland N.A., Abramson R.D., Peters G.J., Petersen G.J., Abramson R.D., Mulahay S.J., Roqueano S.S.,

- RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Sherchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Sodergren E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schatz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Shalska U., Smailus D.E.,  
RA Schnetzer A., Schein J.B., Jones S.J.M., Marras M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RL [6]  
RN RN INTERACTION WITH PDGFBR.  
RX MEDLINE=98114809; PubMed=9742401;  
RA Roche S., Alonso G., Kazlauskas A., Dixit V.M., Courtneyidge S.A.,  
RA Pandey A.;  
RT "Src-like adaptor protein (Slap) is a negative regulator of mitogenesis";  
RT Curr. Biol. 8:975-978 (1998).  
RL [7]  
RN RN CHARACTERIZATION FUNCTION, INTERACTION WITH ZAP70; CD32; VAV1 AND  
RP ICOP, AND MUTAGENESIS OF PRO-72 AND ARG-110.  
RX MEDLINE=20130290; PubMed=1066292;  
RA Sosinowski T., Pandey A., Dixit V.M., Weiss A.;  
RT "Src-like adaptor protein (SLAP) is a negative regulator of T cell receptor signaling";  
RT J. Exp. Med. 191:463-474 (2000).  
RL [8]  
RN RN FUNCTION, SUBCELLULAR LOCALIZATION, MYRISTYLATION, AND MUTAGENESIS OF GLY-1.  
RX MEDLINE=20242033; PubMed=10779329;  
RA Manes G., Bello P., Roche S.;  
RT "Slap negatively regulates Src mitogenic function but does not revert Src-induced cell morphology changes";  
RT Mol. Cell. Biol. 20:3396-3406 (2000).  
RL [9]  
RN RN FUNCTION, AND DEVELOPMENTAL STAGE.  
RX MEDLINE=21453118; PubMed=11567635;  
RA Sosinowski T., Killen N., Weiss A.;  
RT "The Src-like adaptor protein downregulates the T cell receptor on CD4+CD8+ thymocytes and regulates positive selection";  
RT Immunity 15:457-466 (2001).  
CC -!- FUNCTION: Adapter protein, which negatively regulates T-cell receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced activation of nuclear factor of activated T-cells. Involved in the negative regulation of positive selection and mitosis of T-cells. May act by linking signaling proteins such as ZAP70 with CBL, leading to a CBL dependent degradation of signaling proteins.  
CC -!- SUBUNIT: Homodimer. Interacts with phosphorylated CBL, SYK and LAT. Homodimerization and interaction with phosphorylated CBL occurs via its C-terminal domain (By similarity). Interacts with PDPFRS and EPHA2. Interacts with phosphorylated proteins ZAP70; CC -!- SUBCELLULAR LOCATION: Cyttoplasmic; localizes with endosomes.  
CC -!- ALTERNATIVE PRODUCTS:  
Name=1;  
Name=2;  
Name=3;  
IsoId=Q60898-1; Sequence=Displayed;  
IsoId=Q60898-2; Sequence=VSP\_007239;  
-!- TISSUE SPECIFICITY: Predominantly expressed in lymphoid tissues. Highly expressed in spleen, thymus and lymph nodes. Weakly expressed in lung and brain. Expressed in T-cells and at low level in B-cells.  
CC -!- DEVELOPMENTAL STAGE: Expressed during thymocyte maturation. Weakly expressed in CD4 (-) thymocytes, strongly expressed in more CD4 (+) thymocytes, while expression decreases in more mature cells.  
CC -!- DOMAIN: The C-terminal domain is essential for the homodimerization and the interaction with CBL. While the interaction with CBL is apparently mediated via the hydrophobic region of this domain, the highly charged region is apparently

required for the homodimerization (By similarity).

-!- PTM: Phosphorylated (By similarity).

-!- MISCLENOUS: SLA deficient mice show a strong upregulation of TCR and CD5 at the CD4 (+) CD8 (+) stage, and an enhanced positive selection in T-cells.

-!- SIMILARITY: Contains 1 SH2 domain.

-!- SIMILARITY: Contains 1 SH3 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DT 10-OCT-2003 (Rel. 42, Created)

CC DT 10-OCT-2003 (Rel. 42, Last sequence update)

CC DE SRC-like adapter protein (SRC-like adapter protein 1) (hSLAP).

CC GN SLA OR SLAP OR SAPI.

CC OS Homo sapiens (Human)

CC OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CC OC NCBI\_TaxID=9606;

CC CX [1]

RN SEQUENCE FROM N.A., AND INDUCTION BY ATRA.

RP MEDLINE=96423054; PubMed=8825655;

RC TISSUE: Histiocytic lymphoma;

RA Angrist M., Wells D.E., Chakravarti A., Pandey A.,

RA MEDLINE=7144576; PubMed=920066;

RA "Chromosomal localization of the mouse Src-like adapter protein (Slap) gene and its putative human homolog SLA." RT Genomics 30:e23-65 (1995).

RA Miura Y.;

RA "Expression of Src-like adapter protein mRNA is induced by all-trans retinoic acid." RL Biochem. Biophys. Res. Commun. 230:81-84 (1997).

RA [2]

RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RP TISSUE:fetal brain;

RC MEDLINE=98321620; PubMed=9660183;

RA Meijerink P.H.S., Yanakiev P., Zorn I., Griessman A.J., Bikker H.,

RA Dye D., Kalaydjieva L., Baas F.

RA "The gene for the human Src-like adaptor protein (hSLAP) is located within the 64-kb intron of the thyroglobulin gene." RT Eur. J. Biochem. 254:297-303 (1998).

RA [3]

RN SEQUENCE FROM N.A., AND FUNCTION.

RP MEDLINE=1169659; PubMed=1169659;

RA Holland S.J., Liac X.C., Mendenhall M.K., Zhou X., Pardo J., Chu P., Spencer C., Fu A.C., Sheng N., Yu P., Pali E., Nagin A., Shen M., Chan E., Wu X., Li C., Moiseiwitsch M., Aversa G., Koibinger F., Bennett M.K., Molineaux S., Luo Y., Payan D.G., Mancebo H.S.Y., Wu J.;

RA "Functional cloning of Src-like adapter protein-2 (SLAP-2), a novel inhibitor of antigen receptor signaling." RL Eur. J. Exp. Med. 194:1263-1276 (2001).

RA [4]

RN SEQUENCE FROM N.A., AND TISSUE.

RP MEDLINE=2238892; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Coyle F.S., Wagner L., Shenm C.M., Schaefer C.P., Blat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Schaeffer C.P., Blat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer R.D., Rubin G.M., Casavant T.L., Scheetz T.E., Stapleton M., Soares M.B., Bonaldo M.F., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Oranje C., Brownstein M.J., Logellano N.A., Peters G.J., Abramson R.D., Nullahay S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.W., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villeneuve D.K., Muzny D.M., Sodergren E.J., Lu K., Gibbs R.A., Fahey J., Heilman E., Kesteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.,

RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RA [6]

RN SEQUENCE OF 1-71 FROM N.A.

RP MEDLINE=2110045; PubMed=1179692;

RA Kratchmarova I., Sosinowski T., Weiss A., Witter K., Vincenz C., ID SLA-HUMAN STANDARD;

AC Q13239; Q9UMQB;

RA Pandey A.; "Characterization of promoter region and genomic structure of the RT murine and human genes encoding Src like adapter protein.";  
 RL Gene 262:267-273 (2001).  
 [7]

RN ZAP70; CD3Z; SYK AND LAT; WHEN ASSOCIATED WITH S-224 AND S-229.  
 RP FUNCTION: HOMODIMERIZATION, PHOSPHORYLATION, INTERACTION WITH CBL;  
 RP ZAP70; CD3Z; SYK AND LAT; AND MUTAGENESIS OF ARG-110; LEU-217;  
 RP L6U-223; LEU-228 AND 236-LEU-238.  
 RX MEDLINE: 1049770;  
 RA PubMed=1049770;  
 RT "SLAP, a dimeric adapter protein, plays a functional role in T cell receptor signaling.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780(1999).  
 CC -!- FUNCTION: Adapter protein, which negatively regulates T-cell receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced activation of nuclear factor of activated T-cells. Involved in the negative regulation of positive selection and mitosis of T-cells.  
 CC May act by linking signaling proteins such as ZAP70 with CBL, leading to a CBL dependent degradation of signaling proteins.  
 CC -!- SUBUNIT: Interacts with EphB2, VAV1, LCP2 and PDGFRB (By similarity). Homodimer. Homodimerization and interaction with phosphorylated CBL occurs via its C-terminal domain. Interacts with phosphorylated proteins ZAP70, CD3Z, SYK and LAT via its SH2 domain.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; localizes with endosomes (By similarity).  
 CC -!- TISSUE SPECIFICITY: Expressed in lung and fetal brain. Weakly expressed in heart, adult brain, placenta, liver, skeletal muscle, kidney and pancreas.  
 CC -!- INDUCTION: By all-trans retinoic acid (ATRA). Induction is indirect and is mediated through other proteins.  
 CC -!- DOMAIN: The C-terminal domain is essential for the homodimerization and the interaction with CBL. While the interaction with CBL is apparently mediated via the hydrophobic region of this domain, the highly charged region is apparently required for the homodimerization.  
 CC -!- PTM: Phosphorylated.  
 CC -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U30473; AAC50357.1; -. DR EMBL; D89077; BAA13758.1; -. DR EMBL; AAC4403; AAC27662.1; -. DR EMBL; BC000042; AAH07042.1; -. DR EMBL; AU23859.1; CAB53536.1; -. DR HSSP; P08631; 3HCK. DR Genew; HGNC:10902; SLA.  
 DR MIM; 601099; -. DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF00107; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PRO00401; SH2DOMAIN.  
 DR Prodom; PD00093; SH2; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR SH2 domain; SH3 domain; Myristate; Phosphorylation; Lipoprotein.  
 FT INT\_MET 0 BY SIMILARITY.  
 FT DOMAIN 0  
 FT DOMAIN 21 81 SH3.  
 FT DOMAIN 83 174 SH2.  
 FT DOMAIN 211 275 SLA\_C-TERMINAL.  
 FT LIPID 1 1 N-myristoyl glycine (By similarity).  
 FT MUTAGEN 110 110 R->K: STRONGLY REDUCES INTERACTION WITH

ZAP70; CD3Z; SYK AND LAT.  
 L->S: ABOLISHES INTERACTION WITH CBL, WHILE IT DOES NOT AFFECT DIMERIZATION WHEN ASSOCIATED WITH S-224 AND S-229.  
 L->S: ABOLISHES INTERACTION WITH CBL, WHILE IT DOES NOT AFFECT DIMERIZATION WHEN ASSOCIATED WITH S-218 AND S-229.  
 L->S: ABOLISHES INTERACTION WITH CBL, WHILE IT DOES NOT AFFECT DIMERIZATION WHEN ASSOCIATED WITH S-218 AND S-224.  
 L->S: ABOLISHES INTERACTION WITH CBL, WHILE IT DOES NOT AFFECT DIMERIZATION WHEN ASSOCIATED WITH S-218 AND S-224.  
 L->S: OSQ, ABOLISHES INTERACTION WITH CBL, SLIGHTLY AFFECTS DIMERIZATION.

FT MUTAGEN 217 217 ZAP70, CD3Z, SYK AND LAT.  
 FT MUTAGEN 223 223 L->S: ABOLISHES INTERACTION WITH CBL, WHILE IT DOES NOT AFFECT DIMERIZATION WHEN ASSOCIATED WITH S-224 AND S-229.  
 FT MUTAGEN 228 228 L->S: ABOLISHES INTERACTION WITH CBL, WHILE IT DOES NOT AFFECT DIMERIZATION WHEN ASSOCIATED WITH S-218 AND S-229.  
 FT MUTAGEN 236 238 L->S: ABOLISHES INTERACTION WITH CBL, WHILE IT DOES NOT AFFECT DIMERIZATION WHEN ASSOCIATED WITH S-218 AND S-224.  
 FT MUTAGEN 70 70 Y -> D (IN REF. 6).  
 FT CONFLICT 275 AA; 31025 MW; ED683799AD5534F CRC64;  
 SQ SEQUENCE 68 68 Query Match 35.6%; Score 481.5; DB 1; Length 275;  
 Best Local Similarity 40.3%; Pred. No. 4.1e-33;  
 Matches 102; Conservative 43; Missmatches 85; Indels 23; Gaps 4;

Qy 9 KSLPSPSSLSSVQOGCPVTEAERSKATAVLGSPFACGPAAELSLRLGPRLTIVSEDGW 68  
 Db 5 KTPAPA-----ERPLNPPEGDSDFFIAVLSDPSPPISPLPRRGKLVRVSIDEGGW 57

Qy 69 WTVLSEVGREYNIPSVHGVSHGWLWYGLSREKAELLPLPGNPGAFIRESQTRG 128  
 Db 58 WKAISLSTGREGSYIPGICVARYVHGWLFGLGRDKABELLQLPDTKVGFMIRESETRK 117

Qy 129 SYSLSVRLSRPASDRIRHYRHLDNGWLYISPLTPSLQALVDYKHSLEADDICCLIK 188  
 Db 118 FYSLSVR-----HRQVKHYRIFRLPNWYYISRLPQLCDELVNHYSVEADGICCVLT 171

Qy 189 EPCLVQRAGLPGLPDIPLPVQTPLWAKELDSSLFFBEATG----EESLSEL 241  
 Db 172 TPCLUQSTAAPAVASSSSPVTLRQKTVDRVRVSR---LQEDPESTENPTGVDESLSFYCL 228

Qy 242 RESLSFYISNDE 254  
 Db 229 RESIASYSLUTSE 241

RESULT 6  
 LYN\_HUMAN STANDARD; PRT: 511 AA.  
 AC P07948; 129 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 42, Last annotation update)  
 DE Tyrosine-protein kinase LYN (EC 2.7.1.112).  
 GN LYN  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Buterida; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8117210; PubMed=3561390;  
 RA Yamashita Y., Fukuhige S.-I., Sambu K., Sukegawa J., Miyajima N.,  
 RA Matsubara K.-I., Yamamoto T., Toshimura K.;  
 RT "The Yes-related cellular gene lyn encodes a possible tyrosine kinase similar to p56ck.";  
 RL Mol. Cell. Biol. 7:237-243 (1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94171041; PubMed=8125304;  
 RA Rider L.G., Roben N., Miller L., Jelsova C.;  
 RA "The cDNAs encoding two forms of the LYN protein tyrosine kinase are expressed in rat mast cells and human myeloid cells";  
 RL Gene 138:219-222 (1994).  
 CC -!- CATALYTIC ACTIVITY: ATP + Protein tyrosine phosphate.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=LYN A;

Query	66	GDWWTYLSEVGRENTNIPSVHGVK---SHGWLYEGLSREKAELLLELGNGCAFLIR	121
Name-LYN_B;	:	:	
I60D=P07948-2; Sequence=VSP_005002;	:	:	
- - SIMILARITY: Belongs to the Tyr family of protein kinases. SRC subfamily.	:	:	
- - SIMILARITY: Contains 1 SH2 domain.	:	:	
- - SIMILARITY: Contains 1 SH3 domain.	:	:	
EMBL; M16038; AAA55540.1;	:		
EMBL; M79321; AAB50019.1;	:		
PTR; A26719; TVHULY.	:		
HSSP; P04631; IADS.	:		
Genew; HGNC:6735; LYN.	:		
MIM: 165120;	:		
R GO; GO:0004716; F:receptor signaling protein tyrosine kinase . . . ; TAS.	:		
GO; GO:0006468; P:protein amino acid phosphorylation; TAS.	:		
InterPro; IPR0004716; Prot_transduction; TAS.	:		
InterPro; IPR000719; Prot_kinase.	:		
InterPro; IPR001452; SH3.	:		
InterPro; IPR001245; TYR_Pkinase.	:		
InterPro; IPR008266; Tyr_Pkinase_AS.	:		
GO; GO:000165; P:signal transduction; TAS.	:		
InterPro; IPR000719; Prot_kinase.	:		
InterPro; IPR001452; SH3.	:		
InterPro; IPR001245; TYR_Pkinase.	:		
InterPro; IPR008266; Tyr_Pkinase_AS.	:		
Pfam; PF00069; SH2; 1.	:		
Pfam; PF00017; SH2; 1.	:		
Pfam; PF00018; SH3; 1.	:		
PRINTS; PRO0404; SHDOMAIN.	:		
PRINTS; PRO0001; Prot_kinase.	:		
PRINTS; PRO0001; Prot_kinase; 1.	:		
ProDom; PD000003; SH2; 1.	:		
ProDom; PD000056; SH3; 1.	:		
SMART; SM00252; SH2; 1.	:		
SMART; SM0016; SH3; 1.	:		
SMART; SM00219; TyrC; 1.	:		
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	:		
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.	:		
PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.	:		
PROSITE; PS50001; SH2; 1.	:		
PROSITE; PS50002; SH3; 1.	:		
Proto-oncogene; Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding; Myristate; SH2 domain; SH3 domain; INIT_MET	0		
Lipoprotein; Alternative splicing; N-myristoyl glycine (By Similarity).	0		
INIT_NET	0		
LIPID	1		
LIPID	2		
DOM DOMAIN	62	122	
DOM DOMAIN	128	225	
PROTEIN KINASE DOMAIN	246	500	
NP BINDING	252	260	
BINDING	274	274	
ACT-SITE	366	395	
MOD-RES	396	395	
MOD-RES	507	507	
VARSPLIC	22	42	
FT FT			
SEQUENCE	511 AA;	58442 MW;	8419CD461204E364 CRC64;
SQ			
Query Match	27.4%	Score 370.5; DB 1; Length 511;	[4]
Best Local Similarity	39.8%	Prey. No. 1-8e-21;	RP
Matches 80; Conservative	36;	Mismatches 76; Indels 9; Gaps 3;	SEQUENCE OF 163-431 FROM N_A.
6 SRRKSLPSPLSSVQGGPVTMEAERSKATAVGFPAGGAELSLRLGSPLTVSYED	65	RX MEDLINE=0152281; PubMed=2492828	
37 SNQQRPPF-SQLPQRFQTKDPEQGDIVVALLYDGIFPDLSFKGERKMKVLEEH	95	RX "The application of the Polymerase chain reaction to cloning members of the protein tyrosine kinase family.".	
Qy		RT Gene 85:67-74 (1989).	
Db		RT - CATALYTIC ACTIVITY: ATP + a Protein tyrosine = ATP + protein	
		CC - C-terminal phosphate	

-!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=LYN A;  
 CC IsoId=P25911-1; Sequence=Displayed;

CC Name=LYN B;  
 CC IsoId=P25911-2; Sequence=VSP 005003;

CC -!- TISSUE SPECIFICITY: Expressed predominantly in B-lymphoid and  
 myeloid cells.

CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
 -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

DR M6608; AAA39470\_1; -.  
 DR EMBL; M5796; AAA39471\_1; -.  
 DR EMBL; M5797; AAA39472\_1; -.  
 DR EMBL; BC031547; AAU31547\_1; -.  
 DR EMBL; M33426; AAA40017\_1; -.  
 DR HSSP; P08331; IADS.  
 DR MGI; MGI\_96892; LYN.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_Pkinase.  
 DR InterPro; IPR008266; Tyr\_Pkinase\_AS.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PRO00001; Prot\_kinase; 1.  
 DR ProDom; PRO00093; SH2; 1.  
 DR ProDom; PRO00066; SH3; 1.  
 DR SMART; SMART0252; SH2; 1.  
 DR SMART; SMART0326; SH3; 1.  
 DR SMART; SMART0219; Tyr\_kc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 KW Proto-oncogene; Tyrosine\_protein\_kinase; Phosphorylation;  
 KW Transferase; ATP-binding; Myristate; SH2 domain; SH3 domain;  
 KW Palmitate; Lipoprotein; Alternative splicing.

FT INTRIN MET<sub>0</sub>  
 FT LIPID<sub>1</sub>  
 FT LIPID<sub>2</sub>  
 FT DOMAIN<sub>62</sub>  
 FT DOMAIN<sub>128</sub>  
 FT DOMAIN<sub>246</sub>  
 FT NP BIND<sub>252</sub>  
 FT BINDING<sub>274</sub>  
 FT ACT SITE<sub>366</sub>  
 FT MOD RES<sub>396</sub>  
 FT MOD RES<sub>507</sub>  
 FT VAR SPLIC<sub>24</sub>  
 FT CONFLICT<sub>76</sub>  
 FT CONFLICT<sub>160</sub>  
 FT CONFLICT<sub>278</sub>  
 FT CONFLICT<sub>390</sub>  
 FT CONFLICT<sub>414</sub>  
 FT CONFLICT<sub>424</sub>

FT CONFLICT<sub>431</sub>  
 SQ SEQUENCE<sub>511 AA;</sub>  
 Query Match<sub>27.2%</sub>  
 Best Local Similarity<sub>40.3%</sub>  
 Matches<sub>81</sub>; Conservative<sub>33</sub>; Mismatches<sub>78</sub>; Indels<sub>9</sub>; Gaps<sub>3</sub>;

Qy 6 SRRKSLPSPLSSYQQGQPVMTAERSKATAVALGSFAGGPAEFLSRGEPLTIVSD  
 Db 37 SNKQDPPVE-HLPGQRQTKPFECQDIVALVYPDGTHPDDLSPEFKGEKVKVLEH: 95  
 Qy 66 GDWWTYLSEVSGREYNEIPTVHGVKY---SHGWLYEGLSREKABELLIFGPNPGGAFLIR 121  
 Db 96 GEWWRAKSLSSKREGIPNSVAKVNTLEEWPKFDIKRDARQOLLLEGNSAGAFLIR 155  
 Qy 122 ESQTBRGSYSISVRLSPASWDRIBHYRTHCLDNGWLYISPRLFPSQALVDHYSELAD 181  
 Db 156 ESETLKGSFSISVVDQDPMGDKVHKYKIRSLDNGYYISPRIFPCISDMXH1QKQSD 215  
 Qy 182 DICCLKEPCVQLQAGPLPGK 202  
 Db 216 GLCRRLEKACT---SPKEQK 232

RESUME 8  
 LYN RAT STANDARD; PRT; 511 AA.  
 ID LYN RAT Q63320; PRT; 511 AA.  
 AC Q07014; Q63320; PRT; 511 AA.  
 DT 01-JUN-1994 (Rel. 29; Created)  
 DT 01-NOV-1997 (Rel. 35; Last sequence update)  
 DT 10-OCT-2003 (Rel. 42; Last annotation update)  
 DE Tyrosine-protein kinase LYN (EC 2.7.1.112).  
 GN LYN.  
 OS Rattus norvegicus (Rat).  
 OC Mammalia; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID-10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Minoguchi K., Nishikata H., Siraganian R.P.;  
 RA Bacterially expressed rat p55<sup>LN</sup> binds several proteins in rat  
 RA basophilic leukemia cells including pp<sup>62</sup>, a tyrosine phosphorylated  
 RA protein prominent in activated cells.";  
 RL J. Immunol. 150:222-222(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94171041; PubMed-8125304;  
 RA Rider L.G., Raben N., Miller L., Jeltesma C.;  
 RA "The cDNAs encoding two forms of the LYN protein tyrosine kinase are  
 RA expressed in rat mast cells and human myeloid cells.";  
 RL Gene 138:219-222(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97442484; PubMed-9295361;  
 RA Vonakis B.M., Chan H., Haleem-Smith H., Metzger H.;  
 RA "The unique domain at the site on Lyn kinase for its constitutive  
 RA association with the high affinity receptor for IgE.";  
 RA J. Biol. Chem. 272:24072-24080(1997).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC Name=LYN B;  
 CC IsoId=Q07014-2; Sequence=VSP 005004;  
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in B-lymphoid and  
 myeloid cells.  
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi-sib.ch/announce/> or send an email to license@ebi-sib.ch).

CC	DR	EMBL; Li4951; AAA1549_1;	Db	156 ESETLKGSSFLSYRDYDDPMHGDVTKHYKIRSLDNGYYISPRITFPCISDMIKHYQKQSD
CC	DR	EMBL; Li4782; AAA2094_1; -.	Qy	182 DICCLLKEPCVLCQAGPLPGK
CC	DR	EMBL; Li4823; AAA2094_1; -.	Qy	182 DICCLLKEPCVLCQAGPLPGK
CC	DR	EMBL; AF000300; ABT1344_1; -.	Db	216 GLCRRLERACI---SPKPQK
CC	DR	EMBL; AF000301; ABT1345_1; -.		232
CC	DR	EMBL; AF000302; ABT1346_1; -.		
CC	DR	PTR: 156160; I56160.		
CC	DR	HSSP; P0631; IAD5.		
CC	DR	InterPro; IPR00719; Prot_kinase.		
CC	DR	InterPro; IPR00980; SH2.		
CC	DR	InterPro; IPR001452; SH3.		
CC	DR	InterPro; IPR01244; Tyr_pk kinase.		
CC	DR	InterPro; IPR008266; Tyr_pk kinase_AS.		
CC	DR	Pfam; PF00069; pk kinase_1.		
CC	DR	Pfam; PF00017; SH2_1.		
CC	DR	Pfam; PF00018; SH2_1.		
CC	DR	PRINTS; PR00401; SH2DOMAIN.		
CC	DR	PRINTS; PR00422; SH3DOMAIN.		
CC	DR	PRINTS; PR00109; TYRKINASE.		
CC	DR	ProDom; PD000001; Prot kinase; 1.		
CC	DR	ProDom; PD000093; SH2; 1.		
CC	DR	ProDom; PD000066; SH3; 1.		
CC	DR	SMART; SM00252; SH2; 1.		
CC	DR	SMART; SM00326; SH3; 1.		
CC	DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.		
CC	DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.		
CC	DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.		
CC	DR	PROSITE; PS50001; SH; 1.		
CC	DR	PROTO-oncogene Tyrosine-protein kinase; Phosphorylation; Transferase; ARP-binding; Myristate; SH2 domain; SH3 domain; Palmitate; Lipoprotein; Alternative splicing.		
CC	KW	KW		
CC	FT	INIT MET 0 BY SIMILARITY.		
CC	FT	LIPID 1 1 N-myristoyl glycine (By similarity).		
CC	FT	LIPID 2 2 S-palmitoyl cysteine (By similarity).		
CC	FT	DOMAIN 62 122 SH3.		
CC	FT	DOMAIN 128 225 SH2.		
CC	FT	DOMAIN 246 500 PROTEIN KINASE ATP (By similarity).		
CC	FT	NP_BIND 252 260 ATP (By similarity).		
CC	FT	BINDING 274 274 ACT (By similarity).		
CC	FT	ACT_SITE 366 366 BY SIMILARITY.		
CC	FT	MOD_RES 396 396 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).		
CC	FT	MOD_RES 507 507 PHOSPHORYLATION (BY SIMILARITY).		
CC	FT	VARSPLIC 24 44 MISSING (in isoform LYN B).		
CC	FT	PTID_VSP 005004.		
CC	FT	CONFFLICT 230 230 P -> L (IN RBP; 2).		
CC	FT	CONFFLICT 307 307 V -> A (IN REF; 2).		
CC	FT	CONFFLICT 418 418 C -> Y (IN REF; 2).		
CC	FT	SEQUENCE 511 AA; 58529 MW; 24A2E5E229CD43ED CRC64;		
CC	Qy	Query Match Score 367 5; DB 1; Length 511;		
CC	Qy	Best Local Similarity 40.3%; Pred. No. 3.2e-23; Indels 9; Gaps 3;		
CC	Qy	Matches 81; Conservative 33; Mismatches 78;		
CC	Qy	6 SRRKSLDPSPLSLSSVQQGPVTMEARRSKATAVALGSPPAGGPAELSLRLBPLTIVSED 65		
CC	Db	37 SNKQQRVPVE-SQLLPQRFQAKDPEEQGDIIVVALPYDGHDDLSFKRGKVKLEH 95		
CC	Qy	66 GDWTVLSEVSREYNNIPSVHGVKV --- SHGWLYEGLSREKAELLFGNPGQAFILR 121		
CC	Db	96 GEWWKAKLSSREGFPSVNVAYNTLETBWFDPDITKDAAERQQLAGNSAGAFIR 155		
CC	Qy	122 ESCRTRSSESYSLSVRLSPSPASWDRIRYRHCUDNGWLYISPRLLFSPSLQALVDHSELAD 181		
DB	RN	SEQUENCE FROM N.A.		
DB	RN	SEQUENCE OF 21-525 FROM N.A.		
DB	RC	TISSUE-Tissue] mucosa;		
DB	RA	Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Taniguchi A., Fujiiwa T., Ono T., Obayashi M., Nishi T., Ohmori Y., Hirao T., Suzuki Y., Isogai T., Sugano S., "NEBO human cDNA sequencing project.";		
DB	RA	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.		
DB	RL	SEQUENCE FROM N.A.		
DB	RX	MEDLINE=21638749; PubMed=11780052;		

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., RA Mehta M.J., Stavridis G.F., Almeida J.P., Babidge A.K., Bagguley C.L., RA Bailey J., Barlow K.N., Bates K.N., Beard L.M., Bear D.M., RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., RA Buck D., Burriel W.D., Butler A.P., Carder C., Carter N.P., RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Cle C.M., RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., RA Graham M.S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson D., RA Kamber M.P., Kamber G.P., Laird G.K., Lawlor S., RA Leiva-Astiaha M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A., RA Milne S.A., Misery D., Moore M.J.F., Mullikin J.C., Nickerson T., RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., RA Skuse C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E., RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., RA Trotter A., Trotman A.C., Vaudin M., Wall M., Wallis J.M., RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., RA Rogers J., RT "The DNA sequence and comparative analysis of human chromosome 20.", RA RT "The DNA sequence and comparative analysis of human chromosome 20.", RN [6] RN [6]

RP SEQUENCE OF 178-525 FROM N.A.

RC TISSUE=Spleen;

RX MEDLINE=1572549;

RA Hradetzky D., Streibhardt K., Ruegsegger-Waigmann H.;

RT "The genomic locus of the human hemopoietic-specific cell protein tyrosine kinase (PTK)-encoding gene (HCK) confirms conservation of exon-intron structure among human PTKs of the scr family.",; RT Gene 113:275-280(1992).

RL [7]

RP SEQUENCE OF 1-21 FROM N.A., AND ALTERNATIVE INITIATION.

RC MEDLINE=91342636; PubMed=1675927;

RA Lock P., Ralph S., Stanley B., Boulet I., Ramsay R., Dunn A.R.;

RT "Two isoforms of murine hck, generated by utilization of alternative translational initiation codons, exhibit different patterns of tubacellular localization.",; RT Mol. Cell. Biol. 11:4363-4370(1991).

RN [8]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 80-136.

RX MEDLINE=98453315; PubMed=9024658;

RA Sicheri F., Moarefi I., Kurian J.;

RT "Crystal structure of the Src family tyrosine kinase Hck.",; RT Nature 385:602-609 (1997).

RN [9]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 77-525.

RX MEDLINE=9239731; PubMed=9571048;

RA Aroid S., O'Brien R., Frank P., Strub M.P., Hoh F., Dumas C., RA Ladbury J. E.;

RT "RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef.",; RT Biochemistry 37:14683-14691 (1998).

RN [10]

RP STRUCTURE BY NMR OF 77-137.

RX MEDLINE=9263481; PubMed=9109402;

RA Horita D.A., Baldissari D.M., Zhang W., Altieri A.S., Smithgall T.E., RA Gmeiner W.H., Byrd R.A.;

RT "Solution structure of the human Hck SH3 domain and identification of its ligand binding site.",; RT J. Mol. Biol. 278:233-265 (1998).

PL RN [11]

RP STRUCTURE BY NMR OF 138-244.

RX MEDLINE=9109402; PubMed=9109402;

RA Zhang W., Smithgall T.E., Gmeiner W.H., RA "Sequential assignment and secondary structure determination for the SRC homology 2 domain of hematopoietic cellular kinase.",; RT FEBS Lett. 406:131-135 (1997).

CC -!- FUNCTION: May serve as part of a signaling pathway coupling the FC receptor to the activation of the respiratory burst. May also contribute to neutrophil migration and may regulate the granulation process of neutrophils.

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.

CC -!- SUBCELLULAR LOCATION: Isoform p60-HCK and isoform p59-HCK are associated with membranes. Isoform p60-HCK is also cytoplasmic (By similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative initiation; Comment=2 isoforms, p60-HCK (shown here) and p59-HCK, are produced by alternative initiation;

CC TISSUE SPECIFICITY: Expressed predominantly in cells of the myeloid and b-lymphoid lineages.

CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC subfamily. Contains 1 SH2 domain.

CC -!- SIMILARITY: Contains 1 SH3 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; M16591; AAA5643\_1; -

CC DR EMBL; M16592; AAA5644\_1; -

CC DR EMBL; BC014435; AAH14435\_1; -

CC DR EMBL; AK026432; BAB15482\_1; -

CC DR EMBL; AL049539; CAB15560\_1; -

CC DR EMBL; X58741; CAA41565\_2; -

CC DR EMBL; X58742; CAA41565\_2; JOINED.

CC DR EMBL; X58743; CAA41565\_2; JOINED.

CC PIR; A2811; TYRHC.

CC PDB; 2HCK; 20-AUG-97.

CC PDB; 3HCK; 15-JUN-97.

CC PDB; 4HCK; 17-JUN-98.

CC PDB; 5HCK; 17-JUN-98.

CC PDB; 1ADS; 15-MAY-97.

CC PDB; 1BUL; 11-NOV-98.

CC PDB; 1QCF; 08-JUN-99.

CC Genew; HGNC:4840; HCK.

CC MIM: 142370; -

CC GO; GO:0024713; Protein-tyrosine kinase activity; TAS.

CC GO; GO:0007498; P-mesoderm development; TAS.

CC GO; GO:0006468; Protein amino acid phosphorylation; TAS.

CC InterPro; IPR000719; Prot\_kinase.

CC InterPro; IPR000980; SH2.

CC InterPro; IPR001452; SH3.

CC InterPro; IPR001245; TYRKinase.

CC InterPro; IPR008266; Tyr\_Kinase\_AS.

CC Pfam; PF00069; Pfkmase; 1.

CC Pfam; PF000017; SH2; 1.

CC Pfam; PP00018; SH3; 1.

CC PRINTS; PRO0401; SHDOMAIN.

CC PRINTS; PRO0452; SHDOMAIN.

CC PRINTS; PRO00109; TYRKinase.

CC Best Local Similarity 41.6%; Pred 1.3e-22; Mismatches 31; Indels 7; Gaps 2;

CC Matches 77; Conservative 77;

Qy 12 PEPSSLSVQQGPVTMEAKSATAVALSSPAGGPAPSLRQEPLITIVEDGDDWWTV 71

Db 60 PGNSNSNS--NTPGIREAGSEDIVVAVLYDYEATHEDLSFKQDKMYYLEESWWKA 116

Qy 72 LSEVGREYNIPSPVHYGVV---SHGWLYEGLSREKAELILLPGPNPGAFILRESQIRR 127

PROSITE; PS5002; SH3; 1.  
KW Transf erase; Tyrosine\_protein kinase; Phosphorylation; ATP-binding;  
KW Myristate; SH2 domain; SH3 domain; Lipoprotein.  
DR PROSITE; PS5002; SH3; 1.  
KW By SIMILARITY.  
FT INIT-MET 0 0  
FT DOMAIN 55 115  
FT DOMAIN 121 218  
FT DOMAIN 239 492  
FT NP\_BIND 245 253  
FT BINDING 267 267  
FT ACT\_SITE 358 358  
FT LIPID 1 1  
FT MOD\_RES 388 388  
SQ SEQUENCE 503 AA; 56964 MW; B61F3322DDE3436 CRC64;

Query Match 26.5% Score 358; DB 1; Length 503;  
Best Local Similarity 45.3%; Pred. No. 26-22; Gaps 1;  
Matches 72; Conservative 29; Mismatches 54; Indels 4; Gaps 1;

Oy 38 VALGSFPAGGPAAELSLRIGEPLITVSDGDWNTVLSESGREYNIPIAVHGVKV---SHG 93  
Db 61 VALVDYEAIIHHEDLSFQKGDMQWVLEESGENTWKARSLATRKGYIIPSNYVARVDSLETIE 120

Oy 94 WILEGGLESSEKAEPILLIIFGNPGCAFLRERESQRGRSSTLSVRLSRPAWSDRTRHYRTHL 153  
Db 121 WFFKGISKDQAEPQLLAQGNMLGFSWRDSETTKGSYSLSYDPRQGDIVKHYKIRTL 180

Oy 154 DNGMLYISSPRLTPEPSLQLALVHDYSELADDICCLKEPCV 192  
Db 181 DNGGFYIISPRSTFSTLQLBLDHYTKGSISGLCQLSVPCV 219

RESULT 11  
BLK\_HUMAN STANDARD; PRT; 504 AA.  
ID BLK\_HUMAN  
AC Q16281;  
DT 01-OCT-1996 (Rel. 34, Created)  
DE 61 VALDYEAIIHHEDLSFQKGDMQWVLEESGENTWKARSLATRKGYIIPSNYVARVDSLETIE  
GN Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OS Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Macaca.  
OC NCBI\_TaxID=5541;  
OX [1]  
RN RP SEQUENCE FROM N.A.  
RA Picard C., University of Marseille, France.  
CC -!- FUNCTION: May serve as part of a signaling pathway coupling the Fc receptor to the activation of the respiratory burst. May also contribute to neutrophil migration and may regulate the degranulation process of neutrophils (By similarity).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity). Belongs to the Tyr-family of protein kinases. SRC subfamily.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR InterPro: IPRO000719; CR0401.1; -.  
DR InterPro: IPRO00980; SH2.  
DR InterPro: IPRO01451; SH3.  
DR InterPro: IPRO01245; Tyr\_pkinase.  
DR InterPro: IPRO08265; Tyr\_pkinase\_AS.  
DR Pfam: PF00069; pkinase\_1.  
DR Pfam: PF00017; SH2\_1.  
DR Pfam; PF00018; SH3\_1.  
DR PRINTS; PR0041; SH2DOMAIN.  
DR PRINTS; PR00412; SH2DOMAIN.  
DR InterPro: IPRO01245; Tyr\_pkinase.  
DR ProDom; PDD00001; Prot\_kinase; 1.  
DR ProDom; PDD00093; SH2\_1.  
DR SMART; SMART0066; SH3; 1.  
DR SMART; SMART0252; SH3; 1.  
DR SMART; SMART0326; SH3; 1.  
DR SMART; SMART0219; TYRK2; 1.  
DR PROSITE; PS000107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50001; SH2; 1.

RESULT 12  
BLK\_HUMAN STANDARD; PRT; 504 AA.  
ID BLK\_HUMAN  
AC P51451; Q16281;  
DT 01-OCT-1996 (Rel. 34, Created)  
DE 61 VALDYEAIIHHEDLSFQKGDMQWVLEESGENTWKARSLATRKGYIIPSNYVARVDSLETIE  
GN Homo sapiens (Human).  
OC Eukaryote; Metazoa; Chordata; Craniata; Buteleostomi; Homo.  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiridae; Homo.  
NCBI\_TaxID=9606;  
RN SEQUENCE FROM N.A. PubMed=7832795;  
RX MEDLINE=95123078;  
RA Islam K.B.; Rabban H.; Larson C.; Sanders R.; Smith C.I.; RT "Molecular cloning and chromosomal localization of the human homologue of a B-lymphocyte specific protein tyrosine kinase (blk)." ;  
RT "Molecular cloning, characterization, and chromosomal localization of a human lymphoid tyrosine kinase related to murine Blk." ;  
RT J. Immunol. 154:1265-1272(1995).  
RN [2]  
RN SEQUENCE FROM N.A. PubMed=95148218;  
RX Drebin J.A.; Hartzell S.W.; Griffin C.; Campbell M.J.,  
RA Niederhuber J.E.;  
RT "Molecular cloning and chromosomal localization of the human homologue of a B-lymphocyte specific protein tyrosine kinase (blk)." ;  
RT Oncogene 10:477-486(1995).  
CC -!- FUNCTION: May function in a signal transduction pathway that is restricted to B lymphoid cells.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC subfamily.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	DE	Tyrosine-protein kinase HCK (EC 2.7.1.112) (p56-HCK/p59-HCK)
CC	DE	(Hemopoietic cell kinase) (B-cell/myeloid kinase) (BMK).
CC	GN	Mus musculus (Mouse)
CC	OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC	OX	
CC	NCBI_TaxID	10090;
CC	RN	[1]
CC	RP	SEQUENCE OF 21-523 FROM N.A.
CC	RC	STRAIN=ICR; TISSUE=Macrophage;
CC	RX	MEDLINE=88067781; PubMed=3684607;
CC	RA	Klemisz M.J.; McKeercher S.R.; Maki R.A.;
CC	RT	"Nucleotide sequence of the mouse hck gene.";
CC	RL	Nucleic Acids Res 15:9600-9600 (1987).
CC	RN	[2]
CC	RP	SEQUENCE OF 21-523 FROM N.A. AND ALTERNATIVE INITIATION.
CC	RX	MEDLINE=91342636; PubMed=1875927;
CC	RA	Lock P.; Ralph S.; Stanley B.; Boulet I.; Dunn A.R.;
CC	RT	"Two isoforms of murine hck, generated by utilization of alternative translational initiation codons, exhibit different patterns of subcellular localization.";
CC	RT	Mol. Cell. Biol. 11:4353-4370 (1991).
CC	CC	-!- FUNCTION: May serve as part of a signaling pathway coupling the receptor to the activation of the respiratory burst. May also contribute to neutrophil migration, and may regulate the degranulation process of neutrophils.
CC	CC	-!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC	CC	-!- SUBCELLULAR LOCATION: Isoform p59-HCK and isoform p56-HCK are associated with membranes. Isoform p59-HCK is also cytoplasmic.
CC	CC	-!- ALTERNATIVE PRODUCTS:
CC	CC	Comment-2: isoforms, p59-HCK (shown here) and p56-HCK, are produced by alternative initiation;
CC	CC	-!- TISSUE SPECIFICITY: Expressed predominantly in cells of the myeloid and b-lymphoid lineages.
CC	CC	-!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC subfamily.
CC	CC	-!- SIMILARITY: Contains 1 SH2 domain.
CC	CC	-!- SIMILARITY: Contains 1 SH3 domain.
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
CC	EMBL; Y00387; CAA6544_1; -	
CC	PIR; A27232; TMSHAC.	
CC	DR	HSSP; P08631; LAD5.
CC	DR	MGI; MGI:96052; Hck.
CC	DR	InterPro; IPR00108; Neu_Cyt_fact_2.
CC	DR	InterPro; IPR000719; ProC_kinase.
CC	DR	InterPro; IPR000980; SH2.
CC	DR	InterPro; IPR001452; SH3.
CC	DR	InterPro; IPR001245; Tyr_pk kinase.
CC	DR	Pfam; PF00069; kinase_1.
CC	DR	Pfam; PF00017; SH3.
CC	DR	PRINTS; PRD0499; P67PHOX.
CC	DR	PRINTS; PRD0401; SHDDOMAIN.
CC	DR	PRINTS; PRD0452; SH3DOMAIN.
QY	Query Match Score	26.1%; Score 352.5; DB 1; Length 504;
QY	Best Local Similarity	43.6%; Pred. No. 5.7e-22;
QY	Matches	24; Mismatches 68; Indels 5; Gaps 2;
Db	25	PTVMEAEKATAVALGSPAGPAELIRIGEPLTVISEDWDWTUSESGRENTIPS 84
Db	50	PDDEHLDEDKHEVYALYDITATANDRDLQMLKGKLQVLIKGTQDWLRLSLVYREGVPS 109
Qy	85	VHVGVV---HGWLVEGLSREKAEELPFSLQALVYDHYSELADDICCLAREPCV 192
Db	110	NEVARVESLEMMERFFRSQRKEALERQLIAPLKNGPAFLRSQTRGSYSISVRLSRPA 140
Db	110	SWDRIRHYRHICLDNGWLYISPRLTFFPSLQALVYDHYSELADDICCLAREPCV 192
Qy	141	141 AUG-1998 (Rel. 08, Created)
Db	169	169 QGELIKHYKTCRCLDEGGYYISPRITPSLQALVYHYSKKGDSLQRLTLPCV 220
QY	141	Last sequence update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)

RESULT 12  
 HCK\_MOUSE STANDARD PRT: 523 AA.  
 ID HCK\_MOUSE STANDARD  
 AC P08103;  
 DT 01-AUG-1998 (Rel. 08, Created)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)



Query Match Score 351.5; DB 1; Length 502;  
 Best Local Similarity 40.9%; Pprd. No. 6.9e-22;  
 Matches 83; Conservative 32; Mismatches 73; Indels 15; Gaps 4;

DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR000910; SH2.  
 DR InterPro; IPR001422; SH3.  
 DR InterPro; IPR001245; Tyr\_pk kinase.  
 DR InterPro; IPR008266; Tyr\_pk kinase\_AS.  
 DR Pfam; PF00069; Pfam; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PRO0401; SH2DOMAIN.  
 DR PRINTS; PRO0432; SH3DOMAIN.  
 DR PRINTS; PRO1019; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00219; TYKKG; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; FALSE\_NEG.  
 DR PROSITE; PS00001; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 KW Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding;  
 KW Myristate; SH2 domain; SH3 domain; 3D-structure; Lipoprotein.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT LIPID 1 1 N-myristoyl glycine (By similarity).  
 FT DOMAIN 51 111 SH3.  
 FT DOMAIN 117 213 SH2.  
 FT DOMAIN 234 487 PROTEIN\_KINASE.  
 FT DOMAIN 240 248 ATP (BY SIMILARITY).  
 FT BINDING 262 262 ATP (BY SIMILARITY).  
 FT ACT\_SITE 353 353 BY SIMILARITY.  
 FT MOD\_RES 382 382 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT HELIX 124 131  
 FT TURN 132 132  
 FT TURN 134 135  
 FT TURN 138 139  
 FT STRAND 141 145  
 FT TURN 147 148  
 FT STRAND 153 157  
 FT STRAND 159 161  
 FT TURN 162 163  
 FT STRAND 164 166  
 FT STRAND 170 175  
 FT TURN 176 178  
 FT STRAND 179 183  
 FT TURN 184 185  
 FT STRAND 186 188  
 FT HELIX 191 197  
 FT TURN 198 200  
 SQ SEQUENCE 498 AA; 56513 MW; BB49D7B079FDD577 CRC64;  
 Query Match Score 341.5; DB 1; Length 498;  
 Best Local Similarity 40.0%; Pred. No. 4.e-21;  
 Matches 74; Conservative 29; Mismatches 65; Indels 17; Gaps 3;  
 DR 43 PSPN-----QDPDEBEERTVALDYAAANDRLQLVKGERKLQVLRSTDWMLA 90  
 Qy 12 PSPSLSSVQGQGPVTIMEAERSKATAVALGSFPAAGPAEISLRLQEPPLTVSEDGDWWTV 71  
 DR 91 RSLVTRGEGYPSNFVAPVETLEVKWFRTISRDAAQLPANKASFLRESENK 150  
 DR 128 GSYSLSYRLSRPASMDRIRAYRHICLDNGMNYISPLTFPSLQALYDHSELADDICCL 187  
 DR 151 GAFSLSYLK-DITTOGEVVKHYKIRSLDNGGYISPRITPTLQALVQHYSRKGDGLCQKL 209  
 DR 188 KEPCV 192  
 DR 210 TLPCV 214

---

Query Match Score 351.5; DB 1; Length 502;  
 Best Local Similarity 40.9%; Pprd. No. 6.9e-22;  
 Matches 83; Conservative 32; Mismatches 73; Indels 15; Gaps 4;

DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR000910; SH2.  
 DR InterPro; IPR001422; SH3.  
 DR InterPro; IPR001245; Tyr\_pk kinase.  
 DR InterPro; IPR008266; Tyr\_pk kinase\_AS.  
 DR Pfam; PF00069; Pfam; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PRO0401; SH2DOMAIN.  
 DR PRINTS; PRO0432; SH3DOMAIN.  
 DR PRINTS; PRO1019; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00219; TYKKG; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; FALSE\_NEG.  
 DR PROSITE; PS00001; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 KW Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding;  
 KW Myristate; SH2 domain; SH3 domain; 3D-structure; Lipoprotein.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT LIPID 1 1 N-myristoyl glycine (By similarity).  
 FT DOMAIN 51 111 SH3.  
 FT DOMAIN 117 213 SH2.  
 FT DOMAIN 234 487 PROTEIN\_KINASE.  
 FT DOMAIN 240 248 ATP (BY SIMILARITY).  
 FT BINDING 262 262 ATP (BY SIMILARITY).  
 FT ACT\_SITE 353 353 BY SIMILARITY.  
 FT MOD\_RES 382 382 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT HELIX 124 131  
 FT TURN 132 132  
 FT TURN 134 135  
 FT TURN 138 139  
 FT STRAND 141 145  
 FT TURN 147 148  
 FT STRAND 153 157  
 FT STRAND 159 161  
 FT TURN 162 163  
 FT STRAND 164 166  
 FT STRAND 170 175  
 FT TURN 176 178  
 FT STRAND 179 183  
 FT TURN 184 185  
 FT STRAND 186 188  
 FT HELIX 191 197  
 FT TURN 198 200  
 SQ SEQUENCE 498 AA; 56513 MW; BB49D7B079FDD577 CRC64;  
 Query Match Score 341.5; DB 1; Length 498;  
 Best Local Similarity 40.0%; Pred. No. 4.e-21;  
 Matches 74; Conservative 29; Mismatches 65; Indels 17; Gaps 3;  
 DR 43 PSPN-----QDPDEBEERTVALDYAAANDRLQLVKGERKLQVLRSTDWMLA 90  
 Qy 12 LSEVSGREYNIPSVHYGKVS ---HGWLWFGLSREKAELILLPGNPGGAFLRSQTTR 127  
 DR 91 RSLVTRGEGYPSNFVAPVETLEVKWFRTISRDAAQLPANKASFLRESENK 150  
 DR 128 GSYSLSYRLSRPASMDRIRAYRHICLDNGMNYISPLTFPSLQALYDHSELADDICCL 187  
 DR 151 GAFSLSYLK-DITTOGEVVKHYKIRSLDNGGYISPRITPTLQALVQHYSRKGDGLCQKL 209  
 DR 188 KEPCV 192  
 DR 210 TLPCV 214

---

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 tyrosine phosphate. Belongs to the Tyr family of protein kinases. SRC  
 subfamily.

CC -!- SIMILARITY: Contains 1 SH2 domain.

CC -!- SIMILARITY: Contains 1 SH3 domain.

CC -!- SIMILARITY: Contains 1 SH3 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC EMBL; M3003; AAA40453.1; -.

DR PIR; A40092; A40092; -.

DR PDB; 1BLU; 12-MAR-97.

DR MGf; MGI; 88169; Blk.

- RESULT 15
- LCK\_HUMAN STANDARD PRT; 508 AA. Q9N7B8; Q9GDW4; Q9NT8;
- ID LCK\_HUMAN P07100; Q13152; Q9GDW4; Q9NT8;
- AC P06239; P07100; Q13152; Q9GDW4; Q9NT8;
- DT 01-JAN-1988 [Rel. 06, Created]
- DT 01-FEB-1994 [Rel. 28, Last sequence update]
- DT 20-OCT-2003 [Rel. 42, Last annotation update]
- DE Proto-oncogene tyrosine-protein kinase LCK (BC\_2.7.1.112) (P56-LCK)
- DE (LCK) ("T cell-specific protein-tyrosine kinase".)
- DE LCK.
- GN Homo sapiens (Human).
- OS Homo sapiens (Human).
- OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;
- RA Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
- OT NCBI TaxID=9606;
- RN J. Cell. Biochem. 38:117-126(1988).
- [1]
- RN SEQUENCE FROM N.A. MEDLINE=89133626; PubMed=3265417;
- RX Perlmuter R.M., March J.D., Lewis D.B., Peet R., Ziegler S.F.,
- RA Koga Y., Caccia N., Toyonaga B., Spolski R., Yanagi Y., Yoshikai Y.,
- RA Mak T.W.; Will C.B.;
- RT "structure and expression of lck transcripts in human lymphoid
- RT cells."
- RL J. Immunol. 16:1643-1646(1986).
- [2]
- RN SEQUENCE FROM N.A. MEDLINE=90133831; PubMed=3493153;
- RX Perlmuter R.M., March J.D., Lewis D.B., Peet R., Ziegler S.F.,
- RA Koga Y., Caccia N., Toyonaga B., Spolski R., Yanagi Y., Yoshikai Y.,
- RA Mak T.W.; Will C.B.;
- RT "A human T cell-specific cDNA clone (YT16) encodes a protein with extensive homology to a family of protein-tyrosine kinases.";
- RT Eur. J. Immunol. 16:1643-1646(1986).
- [3]
- RN SEQUENCE FROM N.A. MEDLINE=90106967; PubMed=2558056;
- RX Rouer E., van Huygh T., de Souza S.L., Lang M.C., Fischer S.,
- RA Beraudou R.;
- RT "Structure of the human lck gene: differences in genomic organisation within src-related genes affect only N-terminal exons."
- [4]
- RN SEQUENCE FROM N.A. MEDLINE=9105113(1989).
- RX SEQUENCE FROM N.A. VARIANTS LEU-27; GLN-LYS-PRO-231 INS; VAL-352
- RC AND LEU-46, AND PHOSPHORYLATION OF TYR-393 AND TYR-504.
- RT TISSUE=Leukemia;
- RX MEDLINE=91187711; PubMed=8129546;
- RA Wright D.D., Seftor B.M., Kamps M.P.;
- RT "Oncogenic activation of the Lck protein accompanies translocation of the LCK gene in the human HEK2 T-cell leukemia.";
- RT Mol. Cell. Biol. 14:2429-2437(1994).
- RN SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
- RX TISSUE=Leukemic T-cell; MEDLINE=96085119; PubMed=749859;
- RA Vogel L.B., Arthur R., Fujita D.J.,
- RA "An aberrant lck mRNA in two human T-cell lines."
- RA Biochim. Biophys. Acta 1264:168-172(1995).
- [5]
- RN SEQUENCE FROM N.A.
- RC TISSUE=Lymph; MEDLINE=23388257; PubMed=12477932;
- RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuller G.D., Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schaefer C.F., Blatt N.K., Altschul S.F., Zeeberg B., Butow K.H., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiiyuki S., Carninci P., Prange C., Raha S.S., Lognellano N.A., Peters G.J., Abramson R.D., Mullahay S.J., Bosak S.A., McEvran P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Soedergren N., Lu X., Gibbs R.A., Fahey J.U., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
- RN Standard; Q13152; Q9GDW4; Q9NT8;
- AC P06239; P07100; Q13152; Q9GDW4; Q9NT8;
- DT 01-JAN-1988 [Rel. 06, Created]
- DT 01-FEB-1994 [Rel. 28, Last sequence update]
- DT 20-OCT-2003 [Rel. 42, Last annotation update]
- DE Proto-oncogene tyrosine-protein kinase LCK (BC\_2.7.1.112) (P56-LCK)
- DE (LCK) ("T cell-specific protein-tyrosine kinase".)
- DE LCK.
- GN Homo sapiens (Human).
- OS Homo sapiens (Human).
- OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;
- RA Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
- OT NCBI TaxID=9606;
- RN J. Cell. Biochem. 38:117-126(1988).
- [6]
- RN SEQUENCE FROM N.A.
- RC TISSUE=Lymph; MEDLINE=23388257; PubMed=12477932;
- RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuller G.D., Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schaefer C.F., Blatt N.K., Altschul S.F., Zeeberg B., Butow K.H., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., RA Brownstein M.J., Usdin T.B., Toshiiyuki S., Carninci P., Prange C., RA Raha S.S., Lognallano N.A., Peters G.J., Abramson R.D., Mullahay S.J., RA Bosak S.A., McEvran P.J., McKernan K.J., Malek J.A., Gunaratne P.H., RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., RA Villalon D.K., Muzny D.M., Soedergren N., Lu X., Gibbs R.A., RA Fahey J.U., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
- RN Standard; Q13152; Q9GDW4; Q9NT8;
- AC P06239; P07100; Q13152; Q9GDW4; Q9NT8;
- DT 01-JAN-1988 [Rel. 06, Created]
- DT 01-FEB-1994 [Rel. 28, Last sequence update]
- DT 20-OCT-2003 [Rel. 42, Last annotation update]
- DE Proto-oncogene tyrosine-protein kinase LCK (BC\_2.7.1.112) (P56-LCK)
- DE (LCK) ("T cell-specific protein-tyrosine kinase".)
- DE LCK.
- GN Homo sapiens (Human).
- OS Homo sapiens (Human).
- OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;
- RA Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
- OT NCBI TaxID=9606;
- RN J. Cell. Biochem. 38:117-126(1988).
- [7]
- RN SEQUENCE OF 13-508 FROM N.A.
- RC TISSUE=Peripheral blood lymphocytes; MEDLINE=20422621; PubMed=1109057;
- RA Boncristiano M., Majolini M.B., D'Elia M.M., Pacini S., Valensin S., Univieri C., Amedei A., Falini B., Dell Prete G., Telford J.L., Baldari C.T.;
- RT "Defective recruitment and activation of ZAP-70 in common variable immunodeficiency patients with T cell defects."
- RL Eur. J. Immunol. 30:2632-2638(2000).
- [8]
- RN SEQUENCE OF 367-308 FROM N.A.
- RX SEQUENCE OF 374-508 FROM N.A. MEDLINE=87000726; PubMed=389486;
- RA Trevillyan J.M., Lin Y., Chen S.J., Phillips C.A., Canna C., Veilllette A., Foss F.M., Sauvage E.A., Bolen J.B., Rosen N.;
- RA Lima T.J.;
- RT "Expression of the lck tyrosine kinase gene in human colon carcinoma and other non-lymphoid human tumor cell lines."
- RL Oncogene Res. 1:357-374(1987).
- [9]
- RN SEQUENCE OF 367-308 FROM N.A.
- RX SEQUENCE OF 374-508 FROM N.A. MEDLINE=87000726; PubMed=389486;
- RA Garvin A.M., Pawar S., March J.D., Perlmutter R.M.;
- RA "Structure of the murine lck gene and its rearrangement in a murine lymphoma cell line."
- RL Mol. Cell. Biol. 8:3058-3064(1988).
- [10]
- RN SEQUENCE OF 1-34 FROM N.A.
- RX SEQUENCE OF 1-34 FROM N.A. MEDLINE=89096891; PubMed=2850479;
- RA Takadera T., Leung S., Germino A., Koga Y., Takihara Y., Miyamoto N.G., Mak T.W.;
- RA "Structure of the murine lck gene: differential accumulation of two classes of lck transcripts in T cells."
- RL Mol. Cell. Biol. 9:2173-2180(1989).
- [11]
- RN MASS SPECTROMETRY.
- RC TISSUE=Breast cancer; MEDLINE=21829512; PubMed=11840567;
- RA Harris R.A., Yang A., Stein R.C., Lucy K., Brustein L., Herath A., Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J., Zvelebil M.J.;
- RT "Cluster analysis of an extensive human breast cancer cell line protein expression map database."
- RL Proteomics 2:212-223(2002).
- [12]
- RN MASS SPECTROMETRY.
- RC TISSUE=Breast cancer; MEDLINE=94067101; PubMed=7504174;
- RA Vogel L.B., Fujita D.J.;
- RT "The SH3 domain of p56lck is involved in binding to phosphatidylinositol 3'-kinase from T lymphocytes."
- RA Vogel L.B., Fujita D.J.,
- RT "p70 phosphorylates p56lck in T lymphocytes."
- RA Vogel L.B., Fujita D.J.,
- RT "p70 phosphorylates and binds to p56lck is an early event in interleukin-2-induced onset of cell cycle progression in T-lymphocytes."
- RL J. Biol. Chem. 270:2506-2511(1995).
- [15]
- RN PHOSPHORYLATION OF TYR-504.
- RX MEDLINE=92347326; PubMed=1639064;

Berman M., Mustelin T., Oetken C., Partanen J., Flint N.A.,  
 RA Amrein K.E., Autero M., Burn P., Alitalo K.;  
 RT "The human p56ck tyrosine kinase phosphorylates p56ck at Tyr-505 and  
 down regulates its catalytic activity.";  
 RL EMBO J. 11:2919-2924 (1992).  
 [16]  
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 52-225.  
 RX MEDLINE=9203291; PubMed=7512222;  
 RA Eck M.J., Atweel S.K., Shoelson S.E., Harrison S.C.;  
 RT Structure of the regulatory domains of the Src-family tyrosine  
 kinase Lck.;"  
 RT RN Nature 368:764-769 (1994).  
 RL [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 126-220.  
 RX MEDLINE=91173978; PubMed=7532720;  
 RA Mikol V., Baumann G., Keller T.H., Manning U.M., Zurini M.G.M.;  
 RT "The crystal structures of the SH2 domain of p56ck complexed with  
 two phosphopeptides suggest a gated peptide binding site.";  
 RL J. Mol. Biol. 246:344-355 (1995).  
 RN X-RAY CRYSTALLOGRAPHY (1.0 ANGSTROMS) OF 121-225.  
 RX MEDLINE=91177765; PubMed=8604142;  
 RA Tong L., Warren T.C., King J., Betageri R., Ross J., Jakes S.;  
 RT "Crystal structures of the human p56ck SH2 domain in complex with  
 two short phosphotyrosyl peptides at 1.0-A and 1.8-A resolution.";  
 RL J. Mol. Biol. 256:601-610 (1996).  
 RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 118-225.  
 RX MEDLINE=91352059; PubMed=9683372;  
 RA Tong L., Warren T.C., Lukas S., Schembri-King J., Betageri R.,  
 RA Proudfoot J.R., Jakes S.;  
 RT "Carboxymethyl-phenylalanine as a replacement for phosphotyrosine in  
 SH2 domain binding.";  
 RL J. Biol. Chem. 273:20238-20242 (1998).  
 RN [20]  
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 230-500.  
 RX MEDLINE=91100952; PubMed=8945719;  
 RA Yamaguchi H., Hendrickson W.A.;  
 RT "Structural basis for activation of human lymphocyte kinase Lck upon  
 RT tyrosine phosphorylation.";  
 RL Nature 389:484-489 (1996).  
 CC -!- FUNCTION: May participate in antigen-induced T-cell activation.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBUNIT: Binds to phosphatidylinositol 3'-kinase (PI3K) from T  
 CC lymphocytes through its SH3 domain and to the tyrosine  
 CC phosphorylated form of KHDRBS1/P70 through its SH2 domain.  
 CC -!- SUBCELLULAR LOCATION: BOUND TO THE CYTOPLASMIC DOMAIN OF EITHER  
 CC CD4 OR CD8.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=long;  
 CC IsoId=P05239-1; Sequence=Displayed;  
 CC Name=Short;  
 CC IsoId=P02239-2; Sequence=VSP 005000, VSP 005001;  
 CC Note=No experimental confirmation available;  
 CC  
 Query Match Score 340; DB 1; Length 508;  
 Best Local Similarity 40.6%; Fred. No. 6.4e-21;  
 Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
 QY 26 VTMEEAERSKAT-----AVAGSFPPAGGPAAVLRLGEPLTIVSEGDWWTLSEVSQRE 79  
 Db 48 VIVEGSNPPASPLQDNVIAHLYEFSHDGDLGPKGEQLRILEQSGENWKAQSLLTGQE 107  
 QY 80 YNIPSYTHGKVS--HGWLVEGLSREKAELLILPGNPQGAFLIRESQTRGYSLSVR 135  
 Db 108 GRIPFNTVAKANSLEPPWFKNLSKDAEQLAQNTHGFLRESSSTAGESLSVR 167  
 QY 136 ISRPASWDRIHYRHICLDNGMLYISPRITPSIQLAVDHYSELADDICCLKEPCVLR 195  
 Db 168 DFDQNQSEVVHYKIRNLNDNGCFYISPRITPSIQLAVDHYSELADDICCLKEPCVLR 227

Search completed: February 20, 2004, 09:30:30  
 Job time : 41 secs

Result No.	Score	Query Match Length	DB ID	Description
1	370.5	27.4	512	Sequence 16, App1
2	370.5	27.4	512	Sequence 16, App1
3	360.5	26.5	505	Sequence 16, App1
4	360.5	26.5	505	Sequence 17, App1
5	360.5	26.5	505	Sequence 17, App1
6	360.5	26.5	505	Sequence 17, App1
7	344.5	25.5	499	Sequence 19, App1
8	344.5	25.5	499	Sequence 19, App1
9	344.5	25.5	499	Sequence 19, App1
10	340	25.1	508	Sequence 11, App1
11	340	25.1	509	Sequence 17, App1
12	340	25.1	509	Sequence 18, App1
13	340	25.1	509	Sequence 19, App1
14	340	25.1	509	Sequence 19, App1
15	315.5	23.3	537	Sequence 11, App1
16	315.5	23.3	537	Sequence 11, App1
17	315.5	23.3	537	Sequence 11, App1
18	315.5	23.3	537	Sequence 11, App1
19	315.5	23.3	543	Sequence 14, App1
20	315.5	23.3	543	Sequence 14, App1
21	315.5	23.3	543	Sequence 8, App1
22	315.5	23.3	543	Sequence 14, App1
23	313.5	23.2	496	Sequence 2, App1
24	313.5	23.2	496	Sequence 2, App1
25	312.5	23.1	529	Sequence 15, App1
26	312.5	23.1	529	Sequence 15, App1
27	312.5	23.1	529	Sequence 15, App1

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:27:57 ; Search time 45 Seconds  
 (without alignments)  
 299.431 Million cell updates/sec

Title: US-09-939-853A-75  
 Perfect score: 133  
 Sequence: 1 MGSLPSRKSLPSPSLSVSSV.....RESLSFYISLNDEAVSLLDA 261

Scoring table: BLOSUM62  
 GapOp 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued Patents AA:  
 1: /cgtn2\_6/ptodata/2/iaa/SA\_COMBO.PEP:/\*  
 2: /cgtn2\_6/ptodata/2/iaa/5B\_COMBO.DED:/\*  
 3: /cgtn2\_6/ptodata/2/iaa/6A\_COMBO.PEP:/\*  
 4: /cgtn2\_6/ptodata/2/iaa/6B\_COMBO.PEP:/\*  
 5: /cgtn2\_6/ptodata/2/iaa/PCTMUS\_COMBO.PEP:/\*  
 6: /cgtn2\_6/ptodata/2/iaa/backfliest1.pep:/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%  
 Query Match Length DB ID

RESULT 1  
 US-08-426-509A-16  
 Sequence 16, Application US/08426509A  
 Patent No. 6,226469  
 GENERAL INFORMATION:  
 APPLICANT: Ullrich, Axel  
 APPLICANT: Gishinsky, Mikhail  
 APPLICANT: Suresh, Irman G.  
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Penny & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York,  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 1006-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/426,509A  
 FILING DATE: 21-APR-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/232,545  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CORLUZZI, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCUMENT NUMBER: 7683-0074-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 512 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: No. 6326469E  
 US-08-426-509A-16  
 Query Match 27.4%; Score 370.5; DB 4; Length 512;  
 Best Local Similarity 39.8%; Pred. No. 1.7e-31;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

6 SRRKSLBSPSSSSVQGQPTMMEAERSKATAVALGSFPAGGPAAELSLRQEPLTIVSED 65  
 Qy : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 38 SNKQRPVPE-SQLLQQRPFQTQDKPREQDIIVVALPYDGHIPDDLSFKGERMKVLEEH 96  
 66 GDWWTLYSEVGREYNIPSPYHGVKV---SHGWLYEGLSREKAEBELLLPGNPGGAFLIR 121  
 Qy : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 97 GEWKAKSLLTKKEGFI-PSPNTVAKNTLETEEWFFKDITRKDAEROLLAPGNSAGAFLIR 156  
 Qy : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 122 ESOTRRGSSYSVRLSRPASDVRDRYRNGLWYISPRLTPESLQALVDHYSELAD 181  
 Db : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 157 ESETLKSFSLSVRDFDVGDVKHYKIRSILDNGYYISSPRITFFCISDMIKHYQKQAD 216  
 Qy : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 182 DIOCLLKEPCVLRQAGPLPGK 202  
 Db : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 217 GLCRRLBKACI --- SPKPQK 233

RESULT 2  
 US-08-232-545-16  
 Sequence 16, Application US/08232545  
 GENERAL INFORMATION:  
 Patent No. 650578  
 APPLICANT: Ullrich, Axel  
 APPLICANT: Gishizky, Mikhail  
 APPLICANT: Sures, Irman G.  
 TITLE OF INVENTION: No. 6506578 Bel Megakaryocytic Protein Tyrosine Kinases  
 NUMBER OF SEQUENCES: 21  
 ADDRESSSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PacatIn Release #1.0, Version #1.25

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25

ATTORNEY/AGENT INFORMATION:  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/232,545  
 FILING DATE: 22-APR-1994  
 CLASSIFICATION: 435  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-050  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212)790-9090  
 TELEFAX: (212)869-9741  
 TELEX: 66141 PENNIE 16:  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 512 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: protein  
 MOLECULE TYPE: protein

US-08-232-545-16

Query Match 27.4%; Score 370.5; DB 4; Length 512;  
 Best Local Similarity 39.8%; Pred. No. 1.7e-31;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

6 SRRKSLBSPSSSSVQGQPTMMEAERSKATAVALGSFPAGGPAAELSLRQEPLTIVSED 65  
 Qy : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 38 SNKQRPVPE-SQLLQQRPFQTQDKPREQDIIVVALPYDGHIPDDLSFKGERMKVLEEH 96  
 66 GDWWTLYSEVGREYNIPSPYHGVKV---SHGWLYEGLSREKAEBELLLPGNPGGAFLIR 121  
 Qy : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 97 GEWKAKSLLTKKEGFI-PSPNTVAKNTLETEEWFFKDITRKDAEROLLAPGNSAGAFLIR 156

Qy 122 SQTPRGSSVLSRSPASWDRIRHYRIHCLDNGWLYISPRLTFFPSLQLALVDHYSEBLAD 181  
 Db 157 ESETLKGSPLSSTRDFDPVHGDTKHYKIRSLSLDNGGTYISPRITFPCISDMXKHYQKQAD 216

RESULT 4  
 US-08-426-509A-17  
 Sequence 17, Application US/08426509A  
 ; Patent No. 6326669  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ulrich, Axel  
 ; ADDRESS: Pennie & Edmonds  
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: NY  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/426,509A  
 ; FILING DATE: 21-APR-1995  
 ; CLASSIFICATION: 435  
 ; TELECOMMUNICATION INFORMATION:  
 ; PRIORITY APPLICATION NUMBER: 08/232,545  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 863-9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 17:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 505 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-08-426-509A-17

Query Match 26.6%; Score 360.5; DB 4; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30;  
 Sequence 77, Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 Matches 77, Gaps 2;

Qy 12 PSPSLSSSVQGGPVTMEEARSKATAVALGSFPAGGPABSLPGNPGCAFIREQTR 71  
 Db 40 FGPNNSINS---NTPGIREAGSEDIVVALYDYEAIHHELSFKGDQMVLLESGEWKA 96

Query Match 26.6%; Score 360.5; DB 4; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30;  
 Sequence 77, Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 Matches 77, Gaps 2;

Qy 72 LSEVSGREYNIPSPVHGV ---SHGWLVEGLSREKAELLPGNPGCAFIREQTR 71  
 Db 97 RSLATRKEGTYPSPNTVARDSSLETIEWFFKGISKDAEQLDHYSEBLADDICCL 156

Query Match 26.6%; Score 360.5; DB 4; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30;  
 Sequence 77, Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 Matches 77, Gaps 2;

Qy 12 PSPSLSSSVQGGPVTMEEARSKATAVALGSFPAGGPABSLPGNPGCAFIREQTR 71  
 Db 40 FGPNNSINS---NTPGIREAGSEDIVVALYDYEAIHHELSFKGDQMVLLESGEWKA 96

Query Match 26.6%; Score 360.5; DB 4; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30;  
 Sequence 77, Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 Matches 77, Gaps 2;

Qy 72 LSEVSGREYNIPSPVHGV ---SHGWLVEGLSREKAELLPGNPGCAFIREQTR 71  
 Db 97 RSLATRKEGTYPSPNTVARDSSLETIEWFFKGISKDAEQLDHYSEBLADDICCL 156

Query Match 26.6%; Score 360.5; DB 4; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30;  
 Sequence 77, Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 Matches 77, Gaps 2;

Qy 128 GSYSLSVRLSPASNDRIRHYRIHCLDNGWLYISPRLTFFPSLQLALVDHYSEBLADDICCL 187  
 Db 157 GSYSLSVRYDPRQGDTVKHYKIRTLDNGFFYISPRSTSTLQELVDHYKGNGLCQKL 216

RESULT 6  
 PCT-US-05-05008-17  
 Sequence 17, Application PC/TUS9505008  
 ; GENERAL INFORMATION:

APPLICANT: Sugen, Inc.  
 APPLICANT: 515 Galveston Drive  
 Redwood City, California 94063-4720  
 APPLICANT: United States of America  
 APPLICANT: Wissenschaften B.V.  
 Hofgarten Str. 2  
 APPLICANT: München 80539  
 APPLICANT: Germany  
 TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine Kinases

TITLE OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:  
 ADDRESS: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/05008  
 FILING DATE: 24 APR 1995  
 CLASSIFICATION:  
 PCT/US95/05008  
 REFERENCE/DOCKET NUMBER: US 08/232,545  
 FILING DATE: 22 APR 1994  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-074  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 899-9090  
 TELEFAX: (212) 899-9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 17 :  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 505 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 PCT-US95/05008-17

Query Match 26.6%; Score 360 5; DB 5; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30;  
 Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;

Qy 12 PSPSLOSSVQQGPVMEATERSKATAVALGSEPPGSPANLIRGQEPITIVSEGDWTVY 71  
 Db 40 PGPNSHNS--NTPGIREAGSEDIIVVVALDYEAHTHEEDISFGKDQMVLEEGEWKKA 96  
 Qy 72 LSEVSGREYNTIPSVHKGKV---SHOWLYEGLSREKAELLILLPGNPGAFLTRESQPR 127  
 Db 97 RSLATRKEKYPFSNTYARVDSLTEWFFGKISRDAEROLLARQNMGSFMRDSETK 156  
 Qy 128 GSYSLSVYRLSRPASWDRIRHYRIHCILDNGWLYISPRLTFFSLQALNDHYSELADDICCL 187  
 Db 157 GSYSLSVYRLSRPASWDRIRHYRIHCILDNGWLYISPRLTFFSLQALNDHYSELADDICCL 216  
 Qy 188 KEPCV 192  
 Db 217 SVPCM 221

GENERAL INFORMATION:

APPLICANT: Ullrich, Axel  
 APPLICANT: Gishizsky, Mikhail  
 APPLICANT: Sures, Irman G.  
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York,  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: Fast SEQ Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/426,509A  
 FILING DATE: 21-APR-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/232,545  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-0074-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 19 :  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 499 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: No. 6326469  
 US-08-426-509A-19

Query Match 25.5%; Score 344.5; DB 4; Length 499;  
 Best Local Similarity 36.0%; Pred. No. 1.1e-28;  
 Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;

Qy 1 MGSLPSRKSLPSPLSSVQGPV-----TMEAERSK 34  
 Db 1 MGLLSSKRQ----VS EKGKGSWPSNFVAETLE 54  
 Qy 35 ATAVALGSFPGGPABEILSRIGEPLTIVSEGDWTVYSESGRYNTIPSVHKGKV-- 91  
 Db 55 REVVALFDYAVNDRLQLVKLGKLQVLRSVTGKPSNQDPDEE 114  
 Qy 92 -HGWLYIGLSRKAELLILLPGNPGAFLTRESQPRSYSLSRSPASWDRIRHYRI 150  
 Db 115 VKEWFFTRISKDAEQQLAMNKAGSFLRESENNSKAFSLSVK-DITTOGEVVKYKI 173  
 Qy 151 HLDNGWLYISPRLTFFSLQALNDHYSELADDICCLKEPEV 192  
 Db 174 RSLDNGYYISPRITFFTLQALVQHYSKRGDGLCQLTLPCV 215

RESULT 8  
 US-08-426-509A-19  
 Sequence 19, Application US/08426509A  
 ; Patent No. 6326469  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ullrich, Axel  
 ; APPLICANT: Gishizsky, Mikhail  
 ; APPLICANT: Sures, Irman G.  
 ; TITLE OF INVENTION: No. 6326469

RESULT 7  
 US-08-426-509A-19  
 Sequence 19, Application US/08426509A  
 ; Patent No. 6326469

TITLE OF INVENTION: Kinases  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/232,545  
 FILING DATE: 22-APR-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-050  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 869-9090  
 TELEX: (212) 869-9741  
 FAX: (212) 869-9090  
 TELEFAX: (212) 869-9090  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 499 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-08/232-545-19

Query Match 25.5%; Score 344.5; DB 4; Length 499;  
 Best Local Similarity 36.0%; Freq. No. 1.1e-28;  
 Matches 80; Conservative 32; Mismatches 73; Gaps 4;

Qy 1 MGSLPSRRKSPLSPSLSLSSVYQQGPY-----TMEAERSK 34  
 Db 1 MGLLSSKRQ-----VSEKGKGSWPSPKRTQDKAPPPLPVLVFNHLLAPPSPNQDPDEE 54

Qy 35 ATAVALGSFPAGPAAELSLRGAEPLTIVSEGDWTVLSEVSGRYNIPSVHVGKVS-- 91  
 Db 55 REVVALDYAANDRDIQVLGEKLQVLRSTGDNWNLARSWTGREGYVPSNFVAPETEL 114

Qy 92 -HGWLYEGLSREKAELLLPGNPGAAFLIRESQTRGYSYLSTVLSRPAWSDRIRHYRI 150  
 Db 115 VEKWFFRTISKDAEROLLAPKNAKSFLLSEESNGKAFLSVK-DITTOQEVVRYHKI 173

Qy 151 HCLDNGWLYISPLRTEPSLQLVQHYSLEADDICCLIKEPCV 192  
 Db 174 RSLDNGYYISPLRTEPSLQLVQHYSKKGDBLQKLTLPVCV 215

RESULT 9  
 PCT-US95-05008-19  
 GENERAL INFORMATION:  
 APPLICANT: Sugen, Inc.

APPLICANT: 515 Galveston Drive  
 Redwood City, California 94063-4720  
 UNITED STATES OF AMERICA

APPLICANT: Wissenschaften B.V.  
 APPLICANT: Hofgarten Str. 2  
 APPLICANT: München 80539  
 APPLICANT: Germany

TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
 Kinases  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/05008  
 FILING DATE: 24-APR-1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/232,545  
 FILING DATE: 22-APR-1994  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-074  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 869-9090  
 TELEX: (212) 869-9741  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 499 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 PCT/US95-05008-19

Query Match 25.5%; Score 344.5; DB 5; Length 499;  
 Best Local Similarity 36.0%; Freq. No. 1.1e-28;  
 Matches 80; Conservative 32; Mismatches 73; Gaps 4;

Qy 1 MGSLPSRRKSPLSPSLSLSSVYQQGPY-----TMEAERSK 34  
 Db 1 MGLLSSKRQ-----VSEKGKGSWPSPKRTQDKAPPPLPVLVFNHLLAPPSPNQDPDEE 54

Qy 35 ATAVALGSFPAGPAAELSLRGAEPLTIVSEGDWTVLSEVSGRYNIPSVHVGKVS-- 91  
 Db 55 REVVALDYAANDRDIQVLGEKLQVLRSTGDNWNLARSWTGREGYVPSNFVAPETEL 114

Qy 92 -HGWLYEGLSREKAELLLPGNPGAAFLIRESQTRGYSYLSTVLSRPAWSDRIRHYRI 150  
 Db 115 VEKWFFRTISKDAEROLLAPKNAKSFLLSEESNGKAFLSVK-DITTOQEVVRYHKI 173

Qy 151 HCLDNGWLYISPLRTEPSLQLVQHYSLEADDICCLIKEPCV 192  
 Db 174 RSLDNGYYISPLRTEPSLQLVQHYSKKGDBLQKLTLPVCV 215

RESULT 10  
 US-09-862-154-1  
 Sequence 1, Application US/09862154  
 Patent No. 6589758  
 GENERAL INFORMATION:  
 APPLICANT: Zhu, Xidian  
 TITLE OF INVENTION: Crystal of a Kinase-Ligand Complex and Methods of Use  
 FILE REFERENCE: Atty. Docket No. 6589758: A-749  
 CURRENT APPLICATION NUMBER: US/09/862,154  
 NUMBER OF SEQ ID NOS: 1  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 1  
 LENGTH: 508  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-862-154-1

Query Match Score 25.1%; DB 4; Length 508;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

Qy 26 VTMEEAERSKAT----AVALGSFPAGGPAAELSIRLGEPPLIVSEDGDWMTVLSEVSGRE 79  
Db 48 VTYEGSNPPASPLQDNLVIAHLYSYEPSPHDGLGFEKGEOLRLEQGEWWKAQSLSITGQE 107

Qy 80 YNIPSVHGVKS ---HWLVEELSLRSRKAELLLPNPNGGAFLLRESQTRGSLSLR 135  
Db 108 GFIPENFTVAKANSLEPEPWFKNSRDAEROLLAPENTHSFLRESESTAGSFSLSVR 167

Qy 136 LSRPASWDRIRHYRHICLDNGWLYISPRITFPLSQLQALVDHYSELADDCCILKEPVCVLR 195  
Db 169 DFDDQNGEVVKYKIRNLNDGGFYISPRITFGLHELVRYHTNASDGLCTRSLRPCQTQK 228

RESULT 12 US-08-426-509A-18  
Sequence 18, Application US/08426509A  
Patent No. 6326469

GENERAL INFORMATION:  
APPLICANT: Ulrich, Axel  
ADDRESS: Berlin, Germany  
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Penne, Edmunds  
STREET: 1155 Avenue of the Americas  
CITY: New York,  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/426,509A  
FILING DATE: 21-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/232,545  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Crollizzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-0074-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 509 amino acids  
TYPE: amino acid  
STRANDBEADNESS: unknown  
TOPOLOGY: unknown  
US-08-426-509A-18

Query Match Score 25.1%; DB 4; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

Qy 26 VTMEEAERSKAT----AVALGSFPAGGPAAELSIRLGEPPLIVSEDGDWMTVLSEVSGRE 79  
Db 109 GFIPENFTVAKANSLEPEPWFKNSRDAEROLLAPENTHSFLRESESTAGSFSLSVR 168

Qy 49 VTYEGSNPPASPLQDNLVIAHLYSYEPSPHDGLGFEKGEOLRLEQGEWWKAQSLSITGQE 108

Qy 136 LSRPASWDRIRHYRHICLDNGWLYISPRITFPLSQLQALVDHYSELADDCCILKEPVCVLR 195  
Db 169 DFDDQNGEVVKYKIRNLNDGGFYISPRITFGLHELVRYHTNASDGLCTRSLRPCQTQK 228

RESULT 13

US-09-457-040B-8  
 ; Sequence 8, Application US/09457040B  
 ; Patent No. 6387641  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vertex Pharmaceuticals Incorporated  
 ; TITLE OF INVENTION: Crystalized P38 Complexes  
 ; FILE REFERENCE: VPI/9-14  
 ; CURRENT FILING DATE: 1999-12-06  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 8  
 ; LENGTH: 509  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-09-457-040B-8

Query Match 25.1%; Score 340; DB 4; Length 509;  
 Best Local Similarity 40.6%; Fred. No. 3.5e-28;  
 Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

Qy 26 VTMEEAERSKAT----AVALGSFPAGGPAPLRLIGEPLTIVSDGDWTVLSESGRE 79  
 Db 49 VTYEGSNPPASPLQDNVIAHHSYEPHDGFLGEKGEQLRILQESGEWKAQSLTGTQE 108

Qy 80 YNIPSVHYGVKS---HGWLVEGLSREKAEEFLLLPGNPGAAFLFRESCPTGYSLSYR 135  
 Db 109 GTFPENEVAKANSLEPPWFKNLSKDAEFLQLLAGPNTHGFLRESETAGFSLSTR 168

Qy 136 LSRPASWDRIRHYRIHLDNGWLYISPRITPSLQALVDHYSELADDICCLLKSPCIVLQR 195  
 Db 169 DFDQNQGEVVKHYKIRNLNDNGFYISPRITPGLHLVRHYTNASDGLCTRSRCQTQK 228

---

RESULT 15  
 PCT-US95-050008-16  
 ; Sequence 18, Application PC/TU95050008  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sugen, Inc.  
 ; ADDRESS: 515 Galveston Drive  
 ; CITY: Redwood City, California 94063-4720  
 ; STATE: United States of America  
 ; SOCIETY: Wissenschaften E.V.  
 ; TELEPHONE: Hofgarten Str. 2  
 ; FAX: München 80539  
 ; APPLICANT: Germany  
 ; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10016  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/050008  
 ; FILING DATE: 24-APR-1995  
 ; CLASIFICATION INFORMATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30-742  
 ; REFERENCE/DOCKET NUMBER: 7683-074  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEX: (212) 869-9741  
 ; FAX: (212) 869-9741  
 ; INFORMATION FOR SEQ ID NO: 18:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 509 amino acids  
 ; TYPE: amino acid  
 ; STRANDBEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein

---

RESULT 14  
 US-08-232-545-18  
 ; Sequence 18, Application US/08232545  
 ; Patent No. 656578  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ullrich, Axel  
 ; APPLICANT: Gishizky, Mikhail  
 ; APPLICANT: Sures, Irmam G.  
 ; TITLE OF INVENTION: No. 6506378el Megakaryocytic Protein Tyrosine  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10016  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232,545  
 ; FILING DATE: 22-APR-1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30-742  
 ; REFERENCE/DOCKET NUMBER: 7683-050  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 18:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 509 amino acids

PCT-US95-05008-18

Query Match	25.1%	Score 340;	DB 5;	Length 509;
Best Local Similarity	40.6%;	Pred. No. 3	5e-28;	
Matches	73;	Conservative	26;	Mismatches 71; Indels 10; Gaps 2;
Qy	26	VTMBAERSKAT-----AVALGSFPAGGPAELLRLGPPLTIVSDDWWVNLSEVSGRE	79	
Db	49	VTIEGSNPPASPLQDNVIALHSYEPSHDGDLGFEKGEQLRILEQSGEWKAQSLTTGQE	108	
Qy	80	YNTIPSVHYCKVS---HGWLYEGISREXAELLLPGENPGGAFLIRESQTRGSYSLSVR	135	
Db	109	GFIPENFVAKSNSLEPEPMFKNLSRKDAERQILAPGNTGSLIRESESTAGSFSLSVR	168	
Qy	136	LSRPASMDRIRHYTHCLDNGWLYISPRLTFSPLQALVDHYSELADDICCLIKEPCVYLQR	195	
Db	169	DFDQNQGEVTKHYKIRNLNDGGFYISPRTFPGHLHELVRHYTNASDGLCTRUSRPCOTQK	228	

Search completed: February 20, 2004, 09:34:03  
 Job time : 47 secs

ALIGNMENTS									
RESULT 1									
2:	/cgn2_6/_ptodata/2/_pubpaas_PCT_NEW_PUB_pep:*								
3:	/cgn2_6/_ptodata/2/_pubpaas_US06_NEW_PUB_pep:*								
4:	/cgn2_6/_ptodata/2/_pubpaas_US06_PUBCOMB_pep:*								
5:	/cgn2_6/_ptodata/2/_pubpaas_US07_NEW_PUB_pep:*								
6:	/cgn2_6/_ptodata/2/_pubpaas_PCTUS_PUBCOMB_pep:*								
7:	/cgn2_6/_ptodata/2/_pubpaas_US08_NEW_PUB_pep:*								
8:	/cgn2_6/_ptodata/2/_pubpaas_US09A_PUBCOMB_pep:*								
9:	/cgn2_6/_ptodata/2/_pubpaas_US09A_PUBCOMB_pep:*								
10:	/cgn2_6/_ptodata/2/_pubpaas_US09B_PUBCOMB_pep:*								
11:	/cgn2_6/_ptodata/2/_pubpaas_US09C_PUBCOMB_pep:*								
12:	/cgn2_6/_ptodata/2/_pubpaas_US09_NEW_PUB_pep:*								
13:	/cgn2_6/_ptodata/2/_pubpaas_US10A_PUBCOMB_pep:*								
14:	/cgn2_6/_ptodata/2/_pubpaas_US10B_PUBCOMB_pep:*								
15:	/cgn2_6/_ptodata/2/_pubpaas_US10C_PUBCOMB_pep:*								
16:	/cgn2_6/_ptodata/2/_pubpaas_US10_NEW_PUB_pep:*								
17:	/cgn2_6/_ptodata/2/_pubpaas_US60_NEW_PUB_pep:*								
18:	/cgn2_6/_ptodata/2/_pubpaas_US60_PUBCOMB_pep:*								
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
%									
Query Match Length DB ID Description									
No.	Score	Match	Length	DB	ID				
1	1347	99.6	261	14	US-10-043-649-2	Sequence 2, Appli			
2	826	61.0	159	9	US-09-867-550-954	Sequence 954, App			
3	586	43.3	113	9	US-09-867-550-1916	Sequence 1916, Ap			
4	481.5	35.6	276	9	US-09-870-759-64	Sequence 64, Appli			
5	481.5	35.6	276	10	US-09-751-708A-64	Sequence 64, Appli			
6	482.5	35.6	276	14	US-10-043-649-3	Sequence 3, Appli			
7	452.5	33.4	96	9	US-09-867-550-952	Sequence 952, App			
8	370.5	27.4	511	15	US-10-349-322A-42	Sequence 42, Appli			
9	370.5	27.4	512	9	US-09-977-269-16	Sequence 16, Appli			
10	370.5	27.4	512	9	US-09-977-260-16	Sequence 16, Appli			
11	370.5	27.4	512	10	US-09-977-261-16	Sequence 16, Appli			
12	370.5	27.4	512	15	US-10-116-275-162	Sequence 162, Appli			
13	360.5	26.6	505	9	US-09-977-269-17	Sequence 17, Appli			
14	360.5	26.6	505	9	US-09-977-260-17	Sequence 17, Appli			
15	360.5	26.6	505	10	US-09-977-261-17	Sequence 17, Appli			
Query Match Local Similarity 99.6%; Pred. No. 2e-12; Matches 260; Conservative 0; Mismatches 0; Gaps 0;									
RESULT 2									
Sequence 2, Application US-10-043-649-2									
Publication No. US20030059924A1									
GENERAL INFORMATION:									
; APPLICANT: Holland, Sacha J.									
; APPLICANT: Mendenhall, Marcy K.									
; APPLICANT: Pardo, Jorge									
; APPLICANT: Spencer, Collin									
; APPLICANT: Fu, C. Alan									
; APPLICANT: Luo, Ying									
; APPLICANT: Pryan, Donald G.									
; APPLICANT: Mancebo, Helena S.Y.									
; APPLICANT: Wu, Jun									
; APPLICANT: Zhou, Xiulan									
; APPLICANT: Shen, Mary									
; APPLICANT: Liao, X. Charlene									
; APPLICANT: Sheng, Ning									
TITLE OF INVENTION: Cloning of a No. US20030059924A1 Inhibitor of Antigen-receptor-based Functional Screen									
; FILE REFERENCE: A-70219-1.EMS/DHR									
; CURRENT APPLICATION NUMBER: US-10/043-649-2									
; CURRENT FILING DATE: 2002-01-10									
; PRIOR APPLICATION NUMBER: US 60/260,953									
; PRIOR FILING DATE: 2001-01-10									
; NUMBER OF SEQ ID NOS: 3									
; SOFTWARE: Patentin version 3.1									
; SEQ ID NO: 2									
; LENGTH: 261									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
; US-10-043-649-2									
Query Match Local Similarity 99.6%; Pred. No. 2e-12; Matches 260; Conservative 0; Mismatches 0; Gaps 0;									
%									
RESULT 2									
Sequence 2, Application US-10/043-649-2									
Publication No. US 60/260,953									
GENERAL INFORMATION:									
; APPLICANT: Holland, Sacha J.									
; APPLICANT: Mendenhall, Marcy K.									
; APPLICANT: Pardo, Jorge									
; APPLICANT: Spencer, Collin									
; APPLICANT: Fu, C. Alan									
; APPLICANT: Luo, Ying									
; APPLICANT: Pryan, Donald G.									
; APPLICANT: Mancebo, Helena S.Y.									
; APPLICANT: Wu, Jun									
; APPLICANT: Zhou, Xiulan									
; APPLICANT: Shen, Mary									
; APPLICANT: Liao, X. Charlene									
; APPLICANT: Sheng, Ning									
TITLE OF INVENTION: Cloning of a No. US20030059924A1 Inhibitor of Antigen-receptor-based Functional Screen									
; FILE REFERENCE: A-70219-1.EMS/DHR									
; CURRENT APPLICATION NUMBER: US-10/043-649-2									
; CURRENT FILING DATE: 2002-01-10									
; PRIOR APPLICATION NUMBER: US 60/260,953									
; PRIOR FILING DATE: 2001-01-10									
; NUMBER OF SEQ ID NOS: 3									
; SOFTWARE: Patentin version 3.1									
; SEQ ID NO: 2									
; LENGTH: 261									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
; US-10-043-649-2									
%									
RESULT 2									
Sequence 2, Application US-10/043-649-2									
Publication No. US 60/260,953									
GENERAL INFORMATION:									
; APPLICANT: Holland, Sacha J.									
; APPLICANT: Mendenhall, Marcy K.									
; APPLICANT: Pardo, Jorge									
; APPLICANT: Spencer, Collin									
; APPLICANT: Fu, C. Alan									
; APPLICANT: Luo, Ying									
; APPLICANT: Pryan, Donald G.									
; APPLICANT: Mancebo, Helena S.Y.									
; APPLICANT: Wu, Jun									
; APPLICANT: Zhou, Xiulan									
; APPLICANT: Shen, Mary									
; APPLICANT: Liao, X. Charlene									
; APPLICANT: Sheng, Ning									
TITLE OF INVENTION: Cloning of a No. US20030059924A1 Inhibitor of Antigen-receptor-based Functional Screen									
; FILE REFERENCE: A-70219-1.EMS/DHR									
; CURRENT APPLICATION NUMBER: US-10/043-649-2									
; CURRENT FILING DATE: 2002-01-10									
; PRIOR APPLICATION NUMBER: US 60/260,953									

RESULT 2

1 MGSLPSRKSPSPSISVQGGPVTEAERSKATAVALGSFPGGPAELSRLRGELPT 60  
 61 IVSEGDWWTLSVESGRENTIPSVYGVKSHGMWYEGLSREKAELLFGNPGGAFLI 120  
 61 IVSEGDWWTLSVESGRENTIPSVYAKVSHGMWYEGLSREKAELLFGNPGGAFLI 120  
 121 RESQTRGYSISVRLSPASWDRIRHYRHCOLDNGWLYISPRITPSLQLDHYSEL A 180  
 121 RESQTRGYSISVRLSPASWDRIRHYRHCOLDNGWLYISPRITPSLQLDHYSEL A 180  
 121 RESQTRGYSISVRLSPASWDRIRHYRHCOLDNGWLYISPRITPSLQLDHYSEL A 180  
 181 DDICCLKEPCVQLQRAGPLGDIPLPVTQRTPLNKELDSLSEATGEESLSEG 240  
 181 DDICCLKEPCVQLQRAGPLGDIPLPVTQRTPLNKELDSLSEATGEESLSEG 240  
 QY 241 LRESLSFYISLNDAVSLDDA 261  
 Db 241 LRESLSFYISLNDAVSLDDA 261

Query Match Score 43.3%; DB 9; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-50;  
 Matches 0; Indels 0; Gaps 0;

OTHER INFORMATION: Wherein Xaa may be any one of Arg or Gly or Trp  
 SEQ ID NO: 1916  
 LENGTH: 113  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: VARIANT  
 LOCATION: (1)  
 OTHER INFORMATION: Wherein Xaa may be any one of Arg or Gly or Trp  
 US-09-867-550-1916

Query Match Score 586; DB 9; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-50;  
 Matches 0; Indels 0; Gaps 0;

QY 150 IHLCDNGWLYISPLTIPSQALDHYSELADDICCLKEPCVQLQRAGPLGDIPLPVT 209  
 Db 2 IHLCDNGWLYISPLTIPSQALDHYSELADDICCLKEPCVQLQRAGPLGDIPLPVT 61

QY 210 VQTPBLNWKELDSSLILSEATGEESLSEGIRESLSYISLNDAEAVSLDDA 261  
 Db 62 VQTPBLNWKELDSSLILSEATGEESLSEGIRESLSYISLNDAEAVSLDDA 113

RESULT 4

US-09-870-759-64  
 Sequence 64, Application US/09870759  
 Patent No. US2002017751A1

GENERAL INFORMATION:  
 APPLICANT: Leach, Martin D.  
 ATTORNEY OR AGENT NAME: Conley, Pamela Law, Debbie  
 APPLICANT: Topper, James  
 TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and Their Use in the Treatment of Neoplastic Disease

CURRENT APPLICATION NUMBER: US/09/870,759  
 CURRENT FILING DATE: 2000-01-14  
 PRIOR APPLICATION NUMBER: US 60/208,128  
 PRIOR FILING DATE: 2000-05-30  
 NUMBER OF SEQ ID NOS: 2125  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 954  
 LENGTH: 159  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-867-550-954

Query Match Score 61.0%; DB 9; Length 159;  
 Best Local Similarity 99.4%; Pred. No. 2.3e-13;  
 Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSLPSRKSPSPSISVQGGPVTEAERSKATAVALGSFPGGPAELSRLRGELPT 60  
 61 IVSEGDWWTLSVESGRENTIPSVYGVKSHGMWYEGLSREKAELLFGNPGGAFLI 120  
 61 IVSEGDWWTLSVESGRENTIPSVYAKVSHGMWYEGLSREKAELLFGNPGGAFLI 120  
 121 RESQTRGYSISVRLSPASWDRIRHYRHCOLDNGWLY 159  
 Db 121 RESQTRGYSISVRLSPASWDRIRHYRHCOLDNGWLY 159

Query Match Score 35.6%; DB 9; Length 276;  
 Best Local Similarity 40.3%; Pred. No. 4.8e-39;  
 Matches 23; Gaps 4;

QY 9 KSLPSPSISVQGGPVTEAERSKATAVALGSFPGGPAELSRLRGELPTVSBDGW 68  
 6 KSTPAP2 -----ERLPNPEGLDSDFLAVSLDVSPPDISPPIFRGERKLRVISDEGGW 58  
 QY 69 WTVLSESGRENTIPSVYGVKSHGMWYEGLSREKAELLFGNPGGAFLI 120  
 Db 59 WAISLSTGRSEYYPGTCVARYHGMFLFGGRDAEELLQPTKVGSMIRESETKG 118

QY 129 SYSLSYLSRPSMWDRTIRHYRHCOLDNGWLYISPLTIPSQALDHYSELADDICCLKE 188  
 Db 119 FVSLSV-----HRQVHYRIFRLNNWYISPLTQVRLNKELDSSLIFSPRATG----EESLSEGL 172

RESULT 3

US-09-867-550-1916  
 Sequence 1916, Application US/09867550  
 Patent No. US20020082206A1  
 GENERAL INFORMATION:  
 APPLICANT: Leach, Martin D.  
 ATTORNEY OR AGENT NAME: Conley, Pamela Law, Debbie  
 APPLICANT: Topper, James  
 TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and

QY 169 EPCVQLQRAGPLGKDIFLPTVORTPLNKELDSSLIFSPRATG----EESLSEGL 241  
 Db 173 TPCITROSTAAPAVRASSPVTLQRKVTDWRVR-----LQDEDPGTENPLGYDESLFSYGL 229

QY 242 RESLSFYISLNDAEAVSLDDA 254

RESULT 5  
 US-09-751-708A-64  
 / Sequence 64, Application US/09751708A  
 / Publication No. US20030157113A1  
 / GENERAL INFORMATION:  
 / APPLICANT: TEPERMAN, David S  
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
 / CURRENT APPLICATION NUMBER: US/09/751,708A  
 / PRIORITY FILING DATE: 2003-10-15  
 / PRIOR APPLICATION NUMBER: 751708  
 / NUMBER OF SEQ ID NOS: 166  
 / SOFTWARE: Patentin version 3.1  
 / SEQ ID NO: 64  
 / LENGTH: 276  
 / ORGANISM: Homo sapiens  
 / US-09-751-708A-64

Query Match 35.6%; Score 481.5; DB 10; Length 276;  
 Best Local Similarity 40.3%; Pred. No. 4.8e-39;  
 Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;

Qy 9 KSLPSPSSLSSVQCGPVTMEAEERSKATAVALGSPAGCPAELSLRLGPELTIVSEDGDW 68  
 Db 6 KSTPAPA-----ERPLNPBEGLSDFLAVLSDPSPDISPPFRGEKLRVISDEGGW 58

Qy 69 WTVLSEVSGREYNIPSVHVGKVSHGWLYEGLSREXAAELLLPGNPGGAFLIRESQTRG 128  
 Db 59 WKAISLSTGRSYIPICAVRYHWFLEGGRDAEELQLPDTKVGSNIRESETKG 118

Qy 129 SYSLSYLRSRASDRIRHYRHICLDNGNLYISPRLTFSIQLVDHYSELADDICCLIK 188  
 Db 119 FYSLSYR-----HRQYKHTRIFRLPNNTYISPLTFQCLEDLVNHYSEADLGCCVLT 172

Qy 9 KSLPSPSSLSSVQCGPVTMEAEERSKATAVALGSPAGCPAELSLRLGPELTIVSEDGDW 68  
 Db 6 KSTPAPA-----ERPLNPBEGLSDFLAVLSDPSPDISPPFRGEKLRVISDEGGW 58

Qy 69 WTVLSEVSGREYNIPSVHVGKVSHGWLYEGLSREXAAELLLPGNPGGAFLIRESQTRG 128  
 Db 59 WKAISLSTGRSYIPICAVRYHWFLEGGRDAEELQLPDTKVGSNIRESETKG 118

Db 119 FYSLSYR-----HRQYKHTRIFRLPNNTYISPLTFQCLEDLVNHYSEADLGCCVLT 172

Qy 129 SYSLSYLRSRASDRIRHYRHICLDNGNLYISPRLTFSIQLVDHYSELADDICCLIK 188  
 Db 119 FYSLSYR-----HRQYKHTRIFRLPNNTYISPLTFQCLEDLVNHYSEADLGCCVLT 172

Qy 189 EPCVLQDAGPFLPGKDLPLPVTVQRPLNWLKDSSLFSEATRG-----EESLISSEG 241  
 Db 173 TPCLTOSTAAPAVRASSSPPTLROKTVDWRRVSR--LOQDPETGTENPLGVDESFSYGL 229

Qy 242 RESLSFYTSINDE 254  
 Db 230 RESIASLISLSE 242

RESULT 7  
 US-09-867-550-952  
 / Sequence 952, Application US/09867550  
 / Patent No. US20030082206A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Leach, Martin D.  
 / APPLICANT: Mehraban, Fuad,  
 / APPLICANT: Conley, Pamela  
 / APPLICANT: Law, Debbie  
 / APPLICANT: Topper, James  
 / TITLE OF INVENTION: No. US20030082206A1 Polynucleotides from Atherogenic Cells and Their Use  
 / FILE REFERENCE: 214-02-013 (Cura-313)  
 / CURRENT APPLICATION NUMBER: US-09-867-550-952  
 / CURRENT FILING DATE: 2003-09-20  
 / PRIOR APPLICATION NUMBER: US20030082206A1  
 / PRIOR FILING DATE: 2003-05-30  
 / NUMBER OF SEQ ID NOS: 2125  
 / SOFTWARE: FastSEQ for Windows Version 4.0

Qy 1 MGSLPSRKSLPSLSSVQCGGPVTMEAEERSKATAVALGSPAGCPAELSLRLGPELT 60  
 Db 1 MGSLPSRKSLPSLSSVQCGGPVTMEAEERSKATAVALGSPAGCPAELSLRLGPELT 60

RESULT 6  
 US-10-043-649-3  
 / Sequence 3, Application US/10043649  
 / Publication No. US2003005992A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Hollans, Sasha J.  
 / APPLICANT: Mendenhall, Marcy K.  
 / APPLICANT: Pardo, Joyce  
 / APPLICANT: Spencer, Collin  
 / APPLICANT: Fu, C. Alan  
 / APPLICANT: Luo, Ying  
 / APPLICANT: Payan, Donald G.  
 / APPLICANT: Mancebo, Helena S.Y.  
 / APPLICANT: Wu, Jun  
 / APPLICANT: Zhou, Xiuulan  
 / APPLICANT: Shen, Mary  
 / APPLICANT: Liao, X. Charlene  
 / APPLICANT: Sheng, Ning  
 / TITLE OF INVENTION: Cloning of a No. US2003005992A1el Inhibitor of Antigen-receptor  
 / TITLE OF INVENTION: Retrival-based Functional Screen  
 / FILE REFERENCE: A-70219-1/RMS/DHR  
 / CURRENT APPLICATION NUMBER: US/10/C43,649  
 / CURRENT FILING DATE: 2002-01-10  
 / PRIOR APPLICATION NUMBER: US 60/260,953  
 / PRIOR FILING DATE: 2001-01-10  
 / PRIOR FILING DATE: 2001-01-10  
 / PRIOR FILING DATE: 2001-01-10

Qy 61 IVSEGDGMWTVLSEVSGREYNIPSVHVGKVSHGWLYEGLSREXAAELLLPGNPGGAFLI 120  
 Db 61 IVSEGDGMWTVLSEVSGREYNIPSVHVGKVSHGWLYEGLSREXAAELLLPGNPGGAFLI 91

Qy 121 RESET 125  
 Db 92 RESET 96

RESULT 8

; Sequence 42; Application US/10394322A  
; Publication No. US2003022391A1  
; GENERAL INFORMATION:  
; APPLICANT: SONESIS PHARMACEUTICALS, INC.  
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
; FILE REFERENCE: 39750-0006 US  
; CURRENT APPLICATION NUMBER: US/10/394, 322A  
; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: US 60/366, 892  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 42  
; LENGTH: 511  
; TYPE: PRT  
; ORGANISM: Homo sapiens

Query Match 27.4%; Score 370.5; DB 15; Length 511;  
Best Local Similarity 39.8%; Pred. No. 1e-27;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

Db 37 SNKQQRPVPE-SQLLPGQRQTDEEQGDIVVLYPDGTHPDDLSFKKGEMKVLEEH 95

Query 66 GDWWTYLSSESGREYNIPSVHGVK---SHGWLYEGLSREKAELLLPGNPGGAFLIR 121  
Db 96 GEWWKARSLLTKKEFIPSNIVAKINTLEETEWFKDTRDERRQLLAGNSAGAFLIR 155

Query 122 ESQTRGGSYSLSVRLSPASWDRIRHYRTHCLDNGWLYISPRITFPSPQLQALVDHYSELAD 181  
Db 156 ESETLKGSSFLSRSVRDFPVGDVTKHYKRLSDRGYYTSPRITFPCTISDMIKHYQKQAD 215

Query 182 DICCLIKEPCVLRQAGPLPK 202  
Db 216 GLCRRLERAKC---SPKPQK 232

RESULT 9

; Sequence 16; Application US/0997269  
; Patent No. US2002008207A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHZKY, MIKHAIL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1260  
; CURRENT APPLICATION NUMBER: US 09/977, 269  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: US 09/232, 545  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Homo sapiens

Query Match 27.4%; Score 370.5; DB 9; Length 512;  
Best Local Similarity 39.8%; Pred. No. 1e-27;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

Db 38 SNKQQRPVPE-SQLLPGQRQTDEEQGDIVVLYPDGTHPDDLSFKKGEMKVLEEH 95

Query 66 GDWWTYLSSESGREYNIPSVHGVK---SHGWLYEGLSREKAELLLPGNPGGAFLIR 121  
Db 97 GEWWKARSLLTKKEFIPSNIVAKINTLEETEWFKDTRDERRQLLAGNSAGAFLIR 155

Query 122 ESQTRGGSYSLSVRLSPASWDRIRHYRTHCLDNGWLYISPRITFPSPQLQALVDHYSELAD 181  
Db 157 ESETLKGSSFLSRSVRDFPVGDVTKHYKRLSDRGYYTSPRITFPCTISDMIKHYQKQAD 215

Query 182 DICCLIKEPCVLRQAGPLPK 202  
Db 217 GLCRRLERAKC---SPKPQK 232

RESULT 10

; Sequence 16; Application US/09977260  
; Publication No. US0030192790A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHZKY, MIKHAIL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1259  
; CURRENT APPLICATION NUMBER: US/09/977, 261  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232, 545  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1

Query Match 27.4%; Score 370.5; DB 9; Length 512;  
Best Local Similarity 39.8%; Pred. No. 1e-27;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

Db 38 SNKQQRPVPE-SQLLPGQRQTDEEQGDIVVLYPDGTHPDDLSFKKGEMKVLEEH 95

SEQ ID NO 16  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; JWS-09-977-261-16

Query Match Score 370.5; DB 10; Length 512;  
Best Local Similarity 39.8%; Pred. No. 1e-27; Mismatches 76; Indels 9; Gaps 3;

6 SRRKSLPSLSSSSVQGQPVMEARSKATAVALGSFPGGAELSLRLGEPLTIVSDFPAGGAEFLR 65  
38 SINKQORVPVE-SQLLQQRQFTQKDPEEQGDIVVALPYDGHFDLSPPKGEMKVLEH 96

66 GDWNTVLSVSGREYNTPSVHGVY---SHGWLYEGLSREKAELLLLPQNGGAFLIR 121  
97 GEWWAKSLLTKEGIFPSNTVAKNLTEBWFKDITRAEROLLAPGNSAGAFLIR 156

122 ESDTRRGYSLSVRLSRPASMDRIRHYRHCLDNGLWLYSPRATTPSISQUALVHDYSELAD 181  
157 ESETLGKSFSLSVRDFFPVHGDIKYKIRSLDNGYYISPRITEPCISDMIKHYQROAD 216

162 DICCLLKEPCVLRQASPLPGK 202  
217 GLCRRLERKACI----SPKPQK 233

RESULTS 12  
US-10-116-275-162  
; Sequence 162, Application US/10116275  
; Publication No. US20030311476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Darragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116-275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 162  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; JWS-10-116-275-162

Query Match Score 370.5; DB 15; Length 512;  
Best Local Similarity 39.8%; Pred. No. 1e-27; Mismatches 76; Indels 9; Gaps 3;

6 SRRKSLPSLSSSSVQGQPVMEARSKATAVALGSFPGGAELSLRLGEPLTIVSDFPAGGAEFLR 65  
38 SINKQORVPVE-SQLLQQRQFTQKDPEEQGDIVVALPYDGHFDLSPPKGEMKVLEH 96

66 GDWNTVLSVSGREYNTPSVHGVY---SHGWLYEGLSREKAELLLLPQNGGAFLIR 121  
97 GEWWAKSLLTKEGIFPSNTVAKNLTEBWFKDITRAEROLLAPGNSAGAFLIR 156

122 ESDTRRGYSLSVRLSRPASMDRIRHYRHCLDNGLWLYSPRATTPSISQUALVHDYSELAD 181  
157 ESETLGKSFSLSVRDFFPVHGDIKYKIRSLDNGYYISPRITEPCISDMIKHYQROAD 216

182 DICCLLKEPCVLRQASPLPGK 202  
217 GLCRRLERKACI----SPKPQK 233

RESULTS 13

```

US-09-977-269-17 ; Sequence 17, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
;   APPLICANT: GISHIZKY, MIKHAIL
;   APPLICANT: SURES, IRMINGARD
;   TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
;   FILE REFERENCE: 038602/1260
;   CURRENT APPLICATION NUMBER: US/09/977,269
;   CURRENT FILING DATE: 2001-10-16
;   PRIOR FILING DATE: 1994-04-22
;   NUMBER OF SEQ ID NOS: 24
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO: 17
;   LENGTH: 505
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-977-269-17

Query Match Similarity    26.6%; Score 360.5; DB 9; Length 505;
Best Local Similarity 41.6%; Pred. No. 9.6e-27;
Matches 77; Conservative 31; Mismatches 70; Indels 7;
Qy 12 PSPSILSSVQGGPVTMEEAESKATAVALGSFPAGGPAAELSLRLGEPLTIVSDE |
Db 40 PGPNHNS -- NTPGIREAGSEDIVIALYDYEAIHEDLSFQKGDMQMVYLEEBS |

Qy 72 LSEVSGREYNIPSVHYGKV ---SIGHWLYEGLSREKAEEELLPLPQGNPGGAFLIRL |
Db 97 RSLATRKEGYIPSNTYARVDSSLETENWFFGCSRDAERQLLAPGMLGSFMIR |

Qy 128 GSYSLSVRLSRPASWDRIRHYRIHLCLDGNYIISPLTFPSLQALYDHYSELAD |
Db 157 GSYSLSVRLSRPASWDRIRHYRIHLCLDGNYIISPLTFPSLQALYDHYSELAD |

Qy 188 KEPFCV 192
Db 217 SVECM 221

RESULT 14
US-09-977-260-17 ; Sequence 17, Application US/09977260
; GENERAL INFORMATION:
;   APPLICANT: GISHIZKY, MIKHAIL
;   APPLICANT: SURES, IRMINGARD
;   TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
;   FILE REFERENCE: 038602/1160
;   CURRENT APPLICATION NUMBER: US/09/977,260
;   CURRENT FILING DATE: 2001-10-16
;   PRIOR FILING DATE: 1994-04-22
;   NUMBER OF SEQ ID NOS: 24
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO: 17
;   LENGTH: 505
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-977-260-17

Query Match Similarity    26.6%; Score 360.5; DB 9; Length 505;
Best Local Similarity 41.6%; Pred. No. 9.6e-27;
Matches 77; Conservative 31; Mismatches 70; Indels 7;
Qy 12 PSPSILSSVQGGPVTMEEAESKATAVALGSFPAGGPAAELSLRLGEPLTIVSDE |
Db 40 PGPNHNS -- NTPGIREAGSEDIVIALYDYEAIHEDLSFQKGDMQMVYLEEBS |

Qy 72 LSEVSGREYNIPSVHYGKV ---SIGHWLYEGLSREKAEEELLPLPQGNPGGAFLIRL |

```

Db 97 RSLATRKEGYIPSNYVARVDLSLETEEFFKGYSRKDERRQLLAPGMGLGSFMIRDSETK 156  
 Qy 128 GSYSLSVLSPASWDRTRHYRHCLDNGWLYISPRATEPSIQLVTDHYSIADDICCL 187  
 Db 157 GSYSLSVRVDYDPROGDTVKHYKIRTLDNGFYISPRSTSTLQELVDHYKNGNDGLCOKL 216  
 Qy 188 KEPCV 192  
 Db 217 SVPCK 221

---

RESULT 15  
 US-09-977-261-17  
 ; Sequence 17, Application US/09977261  
 ; Publication No. US20030054527A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ULLRICH, AXEL  
 ; APPLICANT: GISHIZKY, MIKHAIL  
 ; APPLICANT: SURES, IRMINGARD  
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
 ; FILE REFERENCE: 039602/1259  
 ; CURRENT APPLICATION NUMBER: US/09/977,261  
 ; CURRENT FILING DATE: 2001-10-16  
 ; PRIOR APPLICATION NUMBER: 08/233,545  
 ; PRIOR FILING DATE: 1994-04-22  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 17  
 ; LENGTH: 505  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ;  
 US-09-977-261-17

Query Match 26 6%; Score 360.5; DB 10; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 9.6e-27;  
 Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 Qy 12 PPSLSSSVQGGPVTMEARRSKATAVALGSEFPGGPABLSRLGEPLTVTSBDGDWTV 71  
 Db 40 PGPNSHNS--NTPGIREAGSEDIVVAYEATHEDLSFGQDQMVVEESGEWKA 96  
 Qy 72 LSEVSREYNIPSPVHGVK--SHGWLYEGLSREKAELLILLPGNPGEAFLRESQTR 127  
 Db 97 RSLATRKEGYIPSNYVARVDLSLETEEFFKGYSRKDERRQLLAPGMGLGSFMIRDSETK 156  
 Qy 128 GSYSLSVLSPASWDRTRHYRHCLDNGWLYISPRATEPSIQLVTDHYSIADDICCL 187  
 Db 157 GSYSLSVRVDYDPROGDTVKHYKIRTLDNGFYISPRSTSTLQELVDHYKNGNDGLCOKL 216  
 Qy 188 KEPCV 192  
 Db 217 SVPCK 221

Search completed: February 20, 2004, 09:50:16  
 Job time : 594 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:18:27 ; Search time 83 Seconds

(without alignments)  
 99.172 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1 MGSILPSRKSILPSPLSSSV.....RESLSFYISLNDEAVSLDDA 261

Sequence: Scoring table: BIOSM62

Gapext 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : SPTREMBL25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mic:\*

8: sp\_organelle:\*

9: sp\_charge:\*

10: sp\_plant:\*

11: sp\_ratod:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_uniclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriab:\*

17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1028	Q7TMX7	259	11	Q7TMX7		Q7tmx7 mus musculus
2	583	Q7TMX7	266	13	Q7TMX7		Q7tmx7 xenopus laevis
3	467.5	Q7TMX7	282	13	Q7TMX7		Q7tmx7 gallus gallus
4	375	Q7TMX7	491	11	Q7TMX7		Q7tmx7 oiceoio mus musculus
5	366.5	Q7TMX7	502	13	Q7TMX7		Q7tmx7 salmo salar
6	361.5	Q7TMX7	488	13	Q7TMX7		Q7tmx7 xenopus laevis
7	352.5	Q7TMX7	505	4	Q7TMX7		Q7tmx7 homo sapiens
8	344.5	Q7TMX7	499	11	Q7TMX7		Q7tmx7 mus musculus
9	340	Q7TMX7	509	6	Q7TMX7		Q7tmx7 hylobates s
10	338	Q7TMX7	509	11	Q7TMX7		Q7tmx7 mus musculus
11	328	Q7TMX7	509	6	Q7TMX7		Q7tmx7 saimiri sciureus
12	320.5	Q7TMX7	519	13	Q7TMX7		Q7tmx7 xenopus laevis
13	315	Q7TMX7	537	11	Q7TMX7		Q7tmx7 rattus norvegicus
14	313.5	Q7TMX7	496	13	Q7TMX7		Q7tmx7 xenopus laevis
15	309.5	Q7TMX7	541	11	Q7TMX7		Q7tmx7 rattus norvegicus
16	309	Q7TMX7	812	15	Q7TMX7		Q7tmx7 avian sarcoma

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SEQUENCES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1028	Q7TMX7	259	11	Q7TMX7		Q7tmx7 mus musculus
2	583	Q7TMX7	266	13	Q7TMX7		Q7tmx7 xenopus laevis
3	467.5	Q7TMX7	282	13	Q7TMX7		Q7tmx7 gallus gallus
4	375	Q7TMX7	491	11	Q7TMX7		Q7tmx7 oiceoio mus musculus
5	366.5	Q7TMX7	502	13	Q7TMX7		Q7tmx7 salmo salar
6	361.5	Q7TMX7	488	13	Q7TMX7		Q7tmx7 xenopus laevis
7	352.5	Q7TMX7	505	4	Q7TMX7		Q7tmx7 homo sapiens
8	344.5	Q7TMX7	499	11	Q7TMX7		Q7tmx7 mus musculus
9	340	Q7TMX7	509	6	Q7TMX7		Q7tmx7 hylobates s
10	338	Q7TMX7	509	11	Q7TMX7		Q7tmx7 mus musculus
11	328	Q7TMX7	509	6	Q7TMX7		Q7tmx7 saimiri sciureus
12	320.5	Q7TMX7	519	13	Q7TMX7		Q7tmx7 xenopus laevis
13	315	Q7TMX7	537	11	Q7TMX7		Q7tmx7 rattus norvegicus
14	313.5	Q7TMX7	496	13	Q7TMX7		Q7tmx7 xenopus laevis
15	309.5	Q7TMX7	541	11	Q7TMX7		Q7tmx7 rattus norvegicus
16	309	Q7TMX7	812	15	Q7TMX7		Q7tmx7 avian sarcoma

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1028	Q7TMX7	259	11	Q7TMX7		Q7tmx7 mus musculus
2	583	Q7TMX7	266	13	Q7TMX7		Q7tmx7 xenopus laevis
3	467.5	Q7TMX7	282	13	Q7TMX7		Q7tmx7 gallus gallus
4	375	Q7TMX7	491	11	Q7TMX7		Q7tmx7 oiceoio mus musculus
5	366.5	Q7TMX7	502	13	Q7TMX7		Q7tmx7 salmo salar
6	361.5	Q7TMX7	488	13	Q7TMX7		Q7tmx7 xenopus laevis
7	352.5	Q7TMX7	505	4	Q7TMX7		Q7tmx7 homo sapiens
8	344.5	Q7TMX7	499	11	Q7TMX7		Q7tmx7 mus musculus
9	340	Q7TMX7	509	6	Q7TMX7		Q7tmx7 hylobates s
10	338	Q7TMX7	509	11	Q7TMX7		Q7tmx7 mus musculus
11	328	Q7TMX7	509	6	Q7TMX7		Q7tmx7 saimiri sciureus
12	320.5	Q7TMX7	519	13	Q7TMX7		Q7tmx7 xenopus laevis
13	315	Q7TMX7	537	11	Q7TMX7		Q7tmx7 rattus norvegicus
14	313.5	Q7TMX7	496	13	Q7TMX7		Q7tmx7 xenopus laevis
15	309.5	Q7TMX7	541	11	Q7TMX7		Q7tmx7 rattus norvegicus
16	309	Q7TMX7	812	15	Q7TMX7		Q7tmx7 avian sarcoma

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1028	Q7TMX7	259	11	Q7TMX7		Q7tmx7 mus musculus
2	583	Q7TMX7	266	13	Q7TMX7		Q7tmx7 xenopus laevis
3	467.5	Q7TMX7	282	13	Q7TMX7		Q7tmx7 gallus gallus
4	375	Q7TMX7	491	11	Q7TMX7		Q7tmx7 oiceoio mus musculus
5	366.5	Q7TMX7	502	13	Q7TMX7		Q7tmx7 salmo salar
6	361.5	Q7TMX7	488	13	Q7TMX7		Q7tmx7 xenopus laevis
7	352.5	Q7TMX7	505	4	Q7TMX7		Q7tmx7 homo sapiens
8	344.5	Q7TMX7	499	11	Q7TMX7		Q7tmx7 mus musculus
9	340	Q7TMX7	509	6	Q7TMX7		Q7tmx7 hylobates s
10	338	Q7TMX7	509	11	Q7TMX7		Q7tmx7 mus musculus
11	328	Q7TMX7	509	6	Q7TMX7		Q7tmx7 saimiri sciureus
12	320.5	Q7TMX7	519	13	Q7TMX7		Q7tmx7 xenopus laevis
13	315	Q7TMX7	537	11	Q7TMX7		Q7tmx7 rattus norvegicus
14	313.5	Q7TMX7	496	13	Q7TMX7		Q7tmx7 xenopus laevis
15	309.5	Q7TMX7	541	11	Q7TMX7		Q7tmx7 rattus norvegicus
16	309	Q7TMX7	812	15	Q7TMX7		Q7tmx7 avian sarcoma

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1028	Q7TMX7	259	11	Q7TMX7		Q7tmx7 mus musculus
2	583	Q7TMX7	266	13	Q7TMX7		Q7tmx7 xenopus laevis
3	467.5	Q7TMX7	282	13	Q7TMX7		Q7tmx7 gallus gallus
4	375	Q7TMX7	491	11	Q7TMX7		Q7tmx7 oiceoio mus musculus
5	366.5	Q7TMX7	502	13	Q7TMX7		Q7tmx7 salmo salar
6	361.5	Q7TMX7	488	13	Q7TMX7		Q7tmx7 xenopus laevis
7	352.5	Q7TMX7	505	4	Q7TMX7		Q7tmx7 homo sapiens
8	344.5	Q7TMX7	499	11	Q7TMX7		Q7tmx7 mus musculus
9	340	Q7TMX7	509	6	Q7TMX7		Q7tmx7 hylobates s
10	338	Q7TMX7	509	11	Q7TMX7		Q7tmx7 mus musculus
11	328	Q7TMX7	509	6	Q7TMX7		Q7tmx7 saimiri sciureus
12	320.5	Q7TMX7	519	13	Q7TMX7		Q7tmx7 xenopus laevis
13	315	Q7TMX7	537	11	Q7TMX7		Q7tmx7 rattus norvegicus
14	313.5	Q7TMX7	496	13	Q7TMX7		Q7tmx7 xenopus laevis
15	309.5	Q7TMX7	541	11	Q7TMX7		Q7tmx7 rattus norvegicus
16	309	Q7TMX7	812	15	Q7TMX7		Q7tmx7 avian sarcoma

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1028	Q7TMX7	259	11	Q7TMX7		Q7tmx7 mus musculus
2	583	Q7TMX7	266	13	Q7TMX7		Q7tmx7 xenopus laevis
3	467.5	Q7TMX7	282	13	Q7TMX7		Q7tmx7 gallus gallus
4	375	Q7TMX7	491	11	Q7TMX7		Q7tmx7 oiceoio mus musculus
5	366.5	Q7TMX7	502	13	Q7TMX7		Q7tmx7 salmo salar
6	361.5	Q7TMX7	488	13	Q7TMX7		Q7tmx7 xenopus laevis
7	352.5	Q7TMX7	505	4	Q7TMX7		Q7tmx7 homo sapiens
8	344.5	Q7TMX7	499	11	Q7TMX7		Q7tmx7 mus musculus
9	340	Q7TMX7	509	6	Q7TMX7		Q7tmx7 hylobates s
10	338	Q7TMX7	509	11	Q7TMX7		Q7tmx7 mus musculus
11	328	Q7TMX7	509	6	Q7TMX7		Q7tmx7 saimiri sciureus
12	320.5	Q7TMX7	519	13	Q7TMX7		Q7tmx7 xenopus laevis
13	315	Q7TMX7	537	11	Q7TMX7		Q7tmx7 rattus norvegicus
14	313.5	Q7TMX7	496	13	Q7TMX7		Q7tmx7 xenopus laevis
15	309.5	Q7TMX7	541	11	Q7TMX7		Q7tmx7 rattus norvegicus
16	309	Q7TMX7	812	15	Q7TMX7		Q7tmx7 avian sarcoma

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1028	Q7TMX7	259	11	Q7TMX7		Q7tmx7 mus musculus
2	583	Q7TMX7	266	13	Q7TMX7		Q7tmx7 xenopus laevis
3	467.5	Q7TMX7	282	13	Q7TMX7		Q7tmx7 gallus gallus
4	375	Q7TMX7</td					

RA	Strausberg R.; Strausberg R.;	Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:1689-1693 (2002).
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; BC05655; AAH5655.1;	
KW	Hypothetical protein.	
SEQUENCE	259 AA;	28495 MW;
Qy	Query Match	Score 1028; DB 11; Length 259;
DR	Best Local Similarity 79.4%; Pred. No. 1.8e-84;	Score 1028; DB 11; Length 259;
DR	Matches 208; Conservative 16; Mismatches 34;	Best Local Similarity 79.4%; Pred. No. 1.8e-84;
DR	EMBL; BCO5635; AAH5635.1; -;	Matches 208; Conservative 16; Mismatches 34;
DR	SEQUENCE 266 AA;	EMBL; BCO5635; AAH5635.1; -;
Qy	1 MGSLPSRKSLPSPSPLSSVQCGQPTMMEARSKATAVALGSPAGGPAAELSLRGEPLT	Query Match 43.1%; Score 583; DB 13; Length 266;
DR	1 MGSLPSRKSLPSPSPLSSVQCGQPTMMEARSKATAVALGSPAGGPAAELSLRGEPLT	Best Local Similarity 47.5%; Pred. No. 2.5e-44;
DR	60 TISEGDGMWTVLSEVSREYNTPSVHYGKVSHGWLYEGLSREKAELLILPGNGGAFLL	Matches 125; Conservative 48; Mismatches 90; Indels 10; Gaps 5;
DR	60 TISEGDGMWTVLSEVSREYNTPSVHYGKVSHGWLYEGLSREKAELLILPGNGGAFLL	Score 583; DB 13; Length 266;
DR	121 RESQTTRGCGSYLSVRLSRPASMDRIRHYRIRHLDNWLWYSPRLLIPSQALVDHYSELLA	Query Match 43.1%; Score 583; DB 13; Length 266;
DR	120 RESQTTRGCGSYLSVRLSRPASMDRIRHYRIRHLDNWLWYSPRLLIPSQALVDHYSELLA	Best Local Similarity 47.5%; Pred. No. 2.5e-44;
DR	181 DDICCLIKRCPVCVQLQAGPLPGKDIPLPVTVQRTPLNKELDSSLLFSEA-ATGEEBSLLE	Matches 125; Conservative 48; Mismatches 90; Indels 10; Gaps 5;
DR	180 DGICCPLRCPVCVQLQGLPKDTPPPVTPVPTISSLNWKKLDRSLLFLEAPSGEASLLE	Score 583; DB 13; Length 266;
DR	240 GLRESLSSYISLNDEAVSLDDA 261	Query Match 43.1%; Score 583; DB 13; Length 266;
DR	240 GLRESLSSYISLNDEAVSLDDA 259	Best Local Similarity 47.5%; Pred. No. 2.5e-44;
RESULT 2		
ID	Q7T0U4	PRELIMINARY; PRT; 266 AA.
AC	Q7T0U4	PRT; 266 AA.
DT	01-OCT-2003 (TREMBrel. 25; Last sequence update)	PRELIMINARY; PRT; 266 AA.
DT	01-OCT-2003 (TREMBrel. 25; Last annotation update)	PRELIMINARY; PRT; 266 AA.
OS	Xenopus laevis (African clawed frog).	RESULT 3
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Q7T3X8
OC	Ampibia; Batrachia; Amura; Mesobatrachia; Pipidae;	AC: Q7T3X8;
XC	Xenopodinae; Xenopus.	DT: 01-OCT-2003 (TREMBrel. 25; Last sequence update)
NCBI_TaxID	[1]	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RN	SEQUENCE FROM N.A.	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RC	TISSUE=Solefin;	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RX	MEDLINE=23341132; PubMed=12454917;	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W., Richardson P.	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative." Dev. Dyn. 225:384-391 (2002).	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RL	SEQUENCE FROM N.A.	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RP	SEQUENCE=Spliced;	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RC	TISSUE=Spliced;	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
RX	Strausberg R.L., Fingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins P.S., Wagner L., Shatten C.M., Schaefer C.P., Bhat N.K., Hopkins R.F., Zeeberg B.R., Butow K.H., Schatz M., Hsieh F., Diatchenko L., Jordan H., Moore T., Max S.I., Wang J., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Logollino N.A., Peters G.J., Abramson R.D., Mullally S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.F., Zeeberg B.R., Hale S., Garie A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Pahey J., Heitton E., Ketteman M., Madan A., Rodriguez S., Sanchez A., Whiting R.W., Maan A.C., Shevchenko Y., Bouffard G.G., Rodriguez A.C., Grinwood J., Schmitz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnierch A., Schein J.E.,	DT: 01-OCT-2003 (TREMBrel. 25; Last annotation update)
Qy	27 TMEABRSKA---TAVA-	Query Match 34.6%; Score 467.5; DB 13; Length 282;
DR	27 TMEABRSKA---TAVA-	Best Local Similarity 41.7%; Pred. No. 7.2e-34;
DR	TMKTPRASAEGGEETIVASSTGQESDFLAVYDPSADISQPPIFHVGKLRVLSDEGWW 63	Matches 105; Conservative 38; Mismatches 38; Indels 33; Gaps 5;
DR	4 TMKTPRASAEGGEETIVASSTGQESDFLAVYDPSADISQPPIFHVGKLRVLSDEGWW 63	Score 467.5; DB 13; Length 282;
DR	70 TVLSEVGRGYNTIPSVHGKVSHGWLYEGLSREKAELLILPGNGGAFIRESOFRGS 129	Query Match 34.6%; Score 467.5; DB 13; Length 282;
DR	64 RVHSLTGRENYIPGVYVAKVYHGNFEGLGREKAEQLPLNTKVGSMPIRESETRKG 123	Best Local Similarity 41.7%; Pred. No. 7.2e-34;
DR	130 YSLSVRLSRPASMDRIRHYRHCLDNGWLYISPRLTTPSILQALVDHYSELADDICCLKE 189	Matches 105; Conservative 38; Mismatches 38; Indels 33; Gaps 5;

RESULT 4		RESULT 5		RESULT 6			
db	124 YSLSRV-----HREYGHYRFLRPNWYTSPRTFOCEDIYHSEADGLUCVLTT 177 Q8CEIO_08CEIO_Preliminary; FRT; 491 AA.	db	190 PCVLDQAGPLGDKDIPILPTVQRTPLNKEI-----DSLLFSEATGSESLSEGR 242 Q8CEIO_08CEIO_Preliminary; FRT; 491 AA.	db	178 PCLTQCTNNHSTIMQVPPVMBKNNWRSHRLMTGDE---STLAAADDSCSYGAR 234 Q8CEIO_08CEIO_Preliminary; FRT; 491 AA.		
dy	243 ESSLFYTISLNDE 254 Q8CEIO_08CEIO_Preliminary; FRT; 491 AA.	dy	235 ESIASIVSILTGD 246 Q8CEIO_08CEIO_Preliminary; FRT; 491 AA.	dy	243 ESSLFYTISLNDE 254 Q8CEIO_08CEIO_Preliminary; FRT; 491 AA.	dy	235 ESIASIVSILTGD 246 Q8CEIO_08CEIO_Preliminary; FRT; 491 AA.
<b>SEQUENCE FROM N.A.</b>							
STRAIN=C57BL/6J; TISSUE=Pancreas; MEDLINE=22354683; PubMed=12466651;							
The RIKEN Genome Exploration Research Group Phase I & II Team, the RIKEN Genome Exploration Research Group Phase I & II Team, "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs"; NATURE 420:563-573 (2002).							
EMBL; AK028112; BAC25753.1; -. NCBI_TaxID=10090; MGI; Mgi:96592; Lyn. GO; GO:0005224; F:ATP binding; IEA. GO; GO:0004574; F:protein serine/threonine kinase activity; IEA. GO; GO:10004713; F:protein-Tyrosine kinase activity; IEA. GO; GO:0007242; P:intracellular signaling cascade; IEA. GO; GO:0005658; P:protein amino acid phosphorylation; IEA. InterPro; IPR000719; Prot_kinase. InterPro; IPR002390; Ser_Thr_pkinase. InterPro; IPR000380; SH2_. InterPro; IPR01152; SH3. InterPro; IPR00145; Tyr_pkinase. InterPro; IPR0008266; Tyr_pkinase_AS. PFAM; PF000089; kinase_1. PFAM; PF00017; SH3; 1. PFAM; PF00018; SH3; 1. PRINTS; PRO00001; SH2DOMAIN. PRINTS; PRO00052; TYRDOMAIN. PRINTS; PRO00009; TYRKINASE. ProDom; PDC000001; Prot_kinase; 1. ProDom; PD000093; SH2; 1. ProDom; PD000066; SH3; 1. SMART; SM00232; SH2; 1. SMART; SM00336; SH3; 1. SMART; SM00230; S_TKC; 1. SMART; SM00239; TyrKC; 1. PROSITE; PS0107; PROTEIN_KINASE_ATP; 1. PROSITE; PS5011; PROTEIN_KINASE_DOM; 1. PROSITE; PS00109; PROTEIN_KINASE_TYR; 1. PROSITE; PS50001; SH2; 1. PROSITE; PS50002; SH3; 1. SEQUENCE 491 AA; 56309 MW; 2C86DB46AA4B6AAB CRC64; SQ							
Query Match 27.7%; Score 375; DB 11; Length 491; Best Local Similarity 39.4%; Pred. No. 3.4e-25; Matches 85; Conservative 35; Mismatches 78; Indels 18; Gaps 4							
1 MGSSLPSRSRK-SLPSPLSSVQ-----GQGPVTPMAERSATAVALGSPPGPAE 50							

Db	1	MGC1KSKRDNLNDDEVDSDKTQPEFHLLPGQRFQTDPEEQGDIVALYPDGIRHDD	60
Qy	51	LSLRLGEPLTIVSBDGDWTTVLSVSGREYNITPSVHGRV---SHGWLYEGLSREKAE	106
Qy	61	LSFPKKGERKVKYLEHGEWTKAKSISSEKREGFISNYAVKNTLTEEFFKDITRKDSER	120
Db	107	LLLPGPNGPAGAFLIRESQCTRGTSLSLSTLSPASWDRHRYTHCLDNGWLTSPIRTF	166
Qy	121	QLLAPGNSAAFLRESETLKGSSLSLSVRYDPMHGDYTKHYKRSLDNGGYYISPIRTF	180
Qy	167	PSIQLALDVHYSSELLADDICCLIKEPCVLRQAGPLPGK	202
Db	181	PCISDMIKHYQKOSPDGLCRRLERAKCI---SPKEQK	212
RESULT 5			
	Q9DDK6	PRELIMINARY;	PRT;
	ID	Q9DDK6	502 AA.
	AC	Q9DDK6;	
	DT	01-MAR-2001	(TRMBLrel_16, Created)
	DT	01-MAR-2001	(TRMBLrel_16, Last sequence update)
	DT	01-OCT-2003	(TRMBLrel_25, Last annotation update)
	DE	Src-Family tyrosine kinase SCK.	
	OS	Salmo salar (Atlantic salmon).	
	OC	Biochemistry; Metabolism; Chordata; Vertebrata; Buteleostei;	
	OC	Actinopterygii; Neoppterygii; Teleostei; Euteleostei;	
	OC	Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.	
	OX	NCBI_Taxid=8030;	
	RN	[1]	
	RP	SEQUENCE FROM N.A.	
	RA	Hordvik I., Male R.;	
	RT	"A leukocyte cDNA of Atlantic salmon encoding a Src-family tyrosine kinase."	
	RL	Submitted (NOV-2000) to the EMBL/CenBank/DBJ databases.	
	CC	-1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.	
	DR	EMBL: AF321110; AAG38611.1; -.	
	DR	HSSPC: P06311; 1ADS.	
	DR	GC: GO:0005524; F:ATP binding; IEA.	
	DR	GO: GO:0004713; F: protein-in-tyrosine kinase activity; IEA.	
	DR	GO: GO:0016740; F: transferase activity; IEA.	
	DR	GO: GO:0007242; P: intracellular signaling cascade; IEA.	
	DR	GO: GO:0004668; P: protein amino acid phosphorylation; IEA.	
	DR	InterPro: IPR000179; Prot_Kinase.	
	DR	InterPro: IPR000980; SH2..	
	DR	InterPro: IPR001452; SH3..	
	DR	InterPro: IPR001245; Tyr_Pkinase..	
	DR	InterPro: IPR008266; Tyr_Pkinase_AS..	
	DR	Pfam: PF00069; Pkinase; 1.	
	DR	Pfam: PF000179; SH2..	
	DR	Pfam: PF00018; SH3..	
	DR	PRINTS: PRO0401; SH2DOMAIN.	
	DR	PRINTS: PRO0452; SH3DOMAIN.	
	DR	PRINTS: PRO009; TYRKINASE.	
	DR	Prodrom: PD000093; Prok_kinase; 1.	
	DR	Prodrom: PD000093; SH2; 1.	
	DR	SMART: SM00052; SH2; 1..	
	DR	SMART: SM00126; SH3; 1..	
	DR	SMART: SM00119; TyrKc..	
	DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.	
	DR	PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.	
	DR	PROSITE: PS00109; PROTEIN_KINASE_Tyr; 1.	
	DR	PROSITE: PS50001; SH3; 1..	
	DR	PROSITE: PS00002; SH3; 1..	
	KW	ATP-binding Kinase; SH3 domain; Transferase; Tyrosine-protein kinase.	
	SQ	SEQUENCE 502 AA;	56600 MW;
Query Match	27.1%	Score 366	DB 13; Length 502;
Best Local Similarity	41.9%	Prev. No. 2,1e-24;	
Matches	80;	Conservative	29; Mismatches 67; Indels 15; Gaps
Qy	15	SISSSVQQGPVPTMEARSKATAVALGSFPGPAAELSLRIGEPLTIVSEDGDNNTVYSE	74

Db 45 STGSPVPSDGP-----ESIAIALYDYEINGEGLGFKKGDKLKLQEGEWRAQSI 96  
 Qy 75 VSGREYNIPSVHYGKVS---HGWLYEGLSREKAEBELLPLPNNFGGAFLIRESQTRGSYS 131  
 Db 97 STGQGFFPSVNTYVADSLTEENFFKGYSRKDEROLLASGMGSTMIRSETTGYS 156  
 Qy 132 LSVRLSRPASWDRTRHYRTHCLONGWLYTSPRLTFPSIQLAQTDVHSELAADDICCLKEPC 191  
 Db 157 LSVRDSQSQGDFTVKHYKRTLDNGGYTSPPRTFTIQELVSHYKKLGDCQALTSPC 216  
 Qy 192 VLQRAGPLPGK 202  
 Db 217 L---SPKPQK 223

RESULT 6  
 O13064 PRELIMINARY; PRT; 488 AA.  
 ID O13064  
 AC O13064;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-OCT-2003 (TREMBlrel. 04, Last sequence update)  
 DE Lyn protein tyrosine kinase.  
 GN LYN.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;  
 OC Xeropoda; Xenopidae; Xenopus.  
 OX NCBITaxonID=8355;  
 RN [1] SEQUENCE FROM N.A.  
 RP RA Fukami Y, Funabiki K., Sato K.;  
 RA "Nucleotide sequence of Xenopus Lyn protein tyrosine kinase.";  
 RT RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; AB003558; BAA20078; 1.  
 DR HSSP; P08631; IAD5.  
 DR GO; GO:000524; F: ATP binding; IEA.  
 DR GO; GO:0004713; P: protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; P: intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P: protein amino acid phosphorylation; IEA.  
 DR GO; GO:0007242; P: intracellular signaling cascade; IBA.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR000719; Prok\_kinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_pkinkase.  
 DR InterPro; IPR000980; SH3.  
 DR PRINTS; PRO0452; SH3DOMAIN.  
 DR Pfam; PF00069; pkinkase\_AS.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PRO0401; SH3; 1.  
 DR DR PRINTS; PRO0452; SH3DOMAIN.  
 DR DR PRINTS; PRO0109; TYRKINASE.  
 DR DR Pfam; PD000001; Prot\_kinase.  
 DR DR Prodrom; PD00009; SH2; 1.  
 DR DR Prodrom; PD00006; SH3; 1.  
 DR DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR DR PROSITE; PS55111; PROTEIN\_KINASE\_DOM; 1.  
 DR DR PROSITE; PS50001; SH2; 1.  
 DR DR PRODom; PD000093; SH2; 1.  
 DR DR SMART; SM0066; SH2; 1.  
 DR DR SMART; SM00326; SH3; 1.  
 DR DR SMART; SM00219; TYRKINASE.  
 DR DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR DR PROSITE; PS50001; SH2; 1.  
 DR DR ATP-binding\_Kinase; SH3; 1. domain: Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 488 AA.; 55794 MW; B7ET06686EA9B2 CRC64;

Query Match 26.1%; Score 361.5; DB 4; Length 505;  
 Best Local Similarity 43.6%; Pred. No. 3.8e-3;  
 Matches 75; Conservative 24; Mismatches 68; Indels 5; Gaps 2;

Db 51 PPDEHEDKHFVVALYDVTAMNDRLQMLKEKLGQVLFKGDFWWLARSLVTVGREGVPS 110  
 Qy 85 VHVGKVS---HGWLYEGLSREKAEBELLPLPNNFGGAFLIRESQTRGSYSVRLSRPA 140  
 DR 111 NPYARVESELEMRWPEFSQGRKEAEROLLAPINKASSFLRESEENKGAFSLVK-DVTT 169

Query Match 26.7%; Score 361.5; DB 13; Length 488;  
 Best Local Similarity 38.3%; Pred. No. 5.6e-24;  
 Matches 82; Conservative 36; Mismatches 79; Indels 17; Gaps 4;

Db 141 SWDRIRRHYRICLNDNGWLYISPLTFFSLQALVDHYSELADDICCLIKEPCV 192  
 Qy

Db 1 MGCIKSKitDNPRDAMGKNNPASRTPSLLPGQ-KMIDQDIEEQGNIVIAIYPQGIGHEDDL 59  
 Qy 53 LRIGEPLITIVSEDGWWTFVLUSENNSGREENNIPSPYHVGKY---SHGMWYEGLSREKAEL 108  
 Db 60 FKKGCEKLVLEEHGEWWKAKSLSTKKECFIPSNTVARYNTLEEEWFKDLTKEAER 119  
 Qy 109 LLPGNPGGFLIRESQTRGSYSVRLSPASWDRHRYHICLDNGWLYISPLTPS 168  
 Db 120 LAPGNNPAGFLIRESETSGKSYSLSTRCDPQGDVLRHYKLTLDNGCYTISPLTPS 179  
 Qy 169 LQALYDHYSBLADDICCLKEPCVTLQRAGPLPGK 202  
 Db 180 INEMIQHYQKQADGLCRKLDKPCF ---SPKPQK 209

RESULT 7  
 Q96INI PRELIMINARY; PRT; 505 AA.  
 ID Q96INI  
 AC Q96INI;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical Protein (B lymphoid tyrosine kinase).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] SEQUENCE FROM N.A.  
 RP TISSUE=Blood, and Lymph;  
 RC Strauberg R.;  
 RA Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; BC007371; AAH07371.1; -.  
 DR EMBL; BC03413; AAH2413.1; -.  
 DR GO; GO:000524; F: ATP binding; IEA.  
 DR GO; GO:0004713; F: protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; P: transferase activity; IEA.  
 DR GO; GO:0007242; P: intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P: protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_pkinkase.  
 DR InterPro; IPR00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR PRINTS; PRO0401; SH3DOMAIN.  
 DR PRINTS; PRO0452; SH3DOMAIN.  
 DR PRINTS; PRO0109; TYRKINASE.  
 DR Prodrom; PD000001; Prot\_kinase.  
 DR Prodrom; PD00009; SH2; 1.  
 DR Prodrom; PD00006; SH3; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS55111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR KW Hypothetical protein, ATP-binding; Kinase; Transferase;  
 KW Tyrosine-protein kinase.  
 SQ SEQUENCE 505 AA; B5F739BEF8389176 CRC64;

Query Match 26.1%; Score 352.5; DB 4; Length 505;  
 Best Local Similarity 43.6%; Pred. No. 3.8e-3;  
 Matches 75; Conservative 24; Mismatches 68; Indels 5; Gaps 2;

Db 25 PTMEEABRSKATAVALGSSFPAGGPANLRLGEPLITIVSEDGDMWTVLSVSGREYNIPS 84  
 Qy 52 PTMEEABRSKATAVALGSSFPAGGPANLRLGEPLITIVSEDGDMWTVLSVSGREYNIPS 84  
 DR 51 PPDEHEDKHFVVALYDVTAMNDRLQMLKEKLGQVLFKGDFWWLARSLVTVGREGVPS 110  
 Qy 85 VHVGKVS---HGWLYEGLSREKAEBELLPLPNNFGGAFLIRESQTRGSYSVRLSRPA 140  
 DR 111 NPYARVESELEMRWPEFSQGRKEAEROLLAPINKASSFLRESEENKGAFSLVK-DVTT 169

Query Match 26.7%; Score 361.5; DB 13; Length 488;  
 Best Local Similarity 38.3%; Pred. No. 5.6e-24;  
 Matches 82; Conservative 36; Mismatches 79; Indels 17; Gaps 4;

Db 141 SWDRIRRHYRICLNDNGWLYISPLTFFSLQALVDHYSELADDICCLIKEPCV 192  
 Qy

Db	RESULT 8 Q8KMB8 ID Q8KMB8 AC Q8KMB8; DT 01-OCT-2002 (TREMBLrel. 22, Created) DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update) DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update) DE B lymphoid kinase. GN BLK. OS Mus musculus (Mouse). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus. [1] RN SEQUENCE FROM N.A. RC TISSUE=Mammary gland; RA Strausberg R.; RT Submitted (MAY-2002) to the EMBL/GenBank/DDJB databases. RN [2] RN SEQUENCE FROM N.A. RC STRAIN=NO; TISSUE=Spleen; RX MEDLINE=22354633; PubMed=12466851; RA THE FANTOM Consortium RA THE RIKEN Genome Exploration Research Group Phase I & II Team; RT "Analysis of the mouse transcriptome based on functional annotation of RT 60,770 full-length cDNAs"; RL Nature 420:563-573 (2002). -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN. CC EMBL; BC00668; AAH0668_1; -; DR EMBL; AK088888; BAC0986_1; -; DR MGD; MGI:88169; DR GO; GO:0004674; F:ATP binding; IEA. DR GO; GO:0004713; F:protein serine/threonine kinase activity; IEA. DR GO; GO:0004740; F:protein tyrosine kinase activity; IEA. DR GO; GO:0007242; F:intracellular signaling cascade; IEA. DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA. DR InterPro; IPR000719; Prot_kinase. DR InterPro; IPR002290; Ser_Thr_Pkinase. DR InterPro; IPR000980; SH2. DR InterPro; IPR001454; SH3. DR InterPro; IPR001245; Tyr_Pkinase. PFam; PF00069; Pkinase; 1. PFam; PF00017; SH2; 1. PFam; PF00018; SH3; 1. DR PRINTS; PR00401; SH2DOMAIN. DR PRINTS; PR00452; SH3DOMAIN. PRINTS; PR00019; TYRKINASE. DR ProDom; PD000093; SH2; 1. DR ProDom; PD000066; SH3; 1. DR SMART; SM00252; SH2; 1. DR SMART; SM00326; SH3; 1. DR SMART; SM00220; S_TKC; 1. DR SMART; SM00219; _TyrKC; 1. DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1. DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1. DR PROSITE; PS50002; SH2; 1. DR PROSITE; PS50001; SH3; 1. KW ATP-binding Kinase; SH3 domain; Transferase; Tyrosine-protein kinase. SEQUENCE 499 AA; 56674 MW; 14C607564BB4R66D CRC64; SQ QY 25.1%; Score 340; DB 6; Length 509; Query Match 25.1%; Score 340; DB 6; Length 509; Best Local Similarity 40.6%; Pred. No. 5.1-21; Matches 73; Conservative 32; Mismatches 73; Indels 37; Gaps 4; 26 VTMEAERSKAT -----AVALGSFPAGGPASLRLGPLELTIVSEDGDWWTVLSEVSRE 79 QY 49 VTYEGSNPPASPLQDNVIALHSYEPHDGLGFGEOLRILEQSGEWWKAQSLTGQE 108 Db 80 YNIPSVHVGVKS -----HGWLYEGLSREKAELLPLPGNFGGAFLLIRESCOTRGSSYSLSPASWDRIRHYRI 135
----	--

Db	109	GFIPNFYAKANSLEPEPWFKNLSRKAERQLLAPGNTGSPFLIRESESTAGSFSLSVR	168	RN [7] SEQUENCE FROM N.A. STRAIN=NOD; TISSUE=Thymus; RC MEDLINE=053013; PubMed=11076861;
Oy	136	LSRPASWDRIRHICLDNGMLYISPLRTFPSLQALVDHYSELADDICCLKEPCVLR	195	RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P., RA Konno H., Akiyama J., Nishi K., Kirsanui T., Tashiro H., Itoh M., RA Sumi N., Ishii Y., Nakamura S., Hachisu M., Ikehagi T., Kasaiwagi K., RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanuki M., RA Yoneda Y., Isikawa T., Ozawa K., Tanaka T., Matsushita S., Kawai J., RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.; RT "RITEN integrated sequence analysis (RISA) system-384 multicapillary sequencer.";
Db	169	DFDQNQEVVKHYKIRNLNDGGFYISPRITPGQHLHYRHTNASDGLCTLSRCPOTQK	228	RT DT 01-OCT-2003 (TREMBLE! 19, Last annotation update) DT 01-OCT-2003 (TREMBLE! 25, Last annotation update) DE Similar to lymphocyte-specific protein tyrosine kinase (2 days neonate thymic cell cDNA, RIKEN full-length enriched library, clone:E30002B06 product:lymphocyte protein tyrosine kinase, full insert sequence). LCK. Mus musculus (Mouse). Metazoal Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus NCBI_TaxID=10090; [1]
GN				OS
GS				OC
DE				OX
DN				OX
GN				NCBI_TaxID=10090; [1]
RP				RN SEQUENCE FROM N.A. RC TISSUE=Salivary gland; RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN				[2] RQ SEQUENCE FROM N.A. RC STRAIN=NOD; TISSUE=Thymus; RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., RA Fukuda S., Furuno M., Hanagata N., Hiramatsu K., Hirakata T., Hirozane T., RA Hayashida K., Hayashi M., Hayashi T., Hara A., Hashizume W., RA Hori F., Imotani K., Ishii Y., Itoh M., Kegawa T., Kasukawa T., RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., RA Saito R., Saitoh H., Sakai C., Sakai K., Sakaue N., Sano H., RA Sasaki D., Shabata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., RA Tagawa A., Takahashi F., Takeki Y., Takeda Y., Tanaka T., RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.; RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN				[3] RQ SEQUENCE FROM N.A. RC STRAIN=NOD; TISSUE=Thymus; RA The RIKEN Genome Exploration Research Group Phase I & II Team, RT "Analysis of the mouse transcriptome based on functional annotation of full-length cDNAs,"; RL Nature 420:1563-1573 (2002).
RP				[4] RQ SEQUENCE FROM N.A. RC STRAIN=NOD; TISSUE=Thymus; RA RIKEN FANTOM Consortium; RT "Functional annotation of a full-length mouse cDNA collection."; RL Nature 409:685-690 (2001).
RN				[5] RQ SEQUENCE FROM N.A. RC STRAIN=NOD; TISSUE=Thymus; RX MEDLINE=20499374; PubMed=11042159; RA Carninci P., Shibata Y., Sugahara Y., Shibata K., Itoh M., RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.; RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."; RL Genome Res. 10:1617-1630 (2000).
RP				[6] RQ SEQUENCE FROM N.A. RC STRAIN=NOD; TISSUE=Thymus; RX MEDLINE=99279253; PubMed=10349636; RA Carninci P., Hayashizaki Y.; RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.; RT "High-efficiency full-length cDNA cloning."; RL Meth. Enzymol. 303:19-44 (1999).
RP				[7] RQ SEQUENCE FROM N.A. RC MEDLINE=20499374; PubMed=11042159; RA Carninci P., Shibata Y., Sugahara Y., Shibata K., Itoh M., RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.; RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."; RL Genome Res. 10:1617-1630 (2000).

RESULT 12	
Q1-DEC-2001 (TREMBLrel. 19, Created)	PRT; 519 AA.
01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
Lck tyrosine kinase (Fragment).	
Saimiri sciureus (Common squirrel monkey):	
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Placentalia; Cebidae; Cebinae; Saimiri.	
NCBI TaxID:9521;	
[1]	
SEQUENCE FROM N.A.	
MEDLINE=1424508; PubMed=1533187;	
Greve T, Tamgoune G., Fleischner B., Fickenscher H., Broeker B.M.;	
"Downregulation of p56 <sup>ck</sup> tyrosine kinase activity in T cells of squirrel monkeys (Saimiri sciureus) correlates with the non-transforming and apathogenic properties of herpesvirus saimiri in its natural host";	
J. Virol. 75: 9252-9261(2001).	
-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.	
EMBL; AJ777921; CRAC38B71.1; -.	
GO; GO:0005524; F-ATP binding; IEA.	
GO; GO:0001740; F-tyrosine kinase activity; IEA.	
GO; GO:0007242; Protein intracellular signaling cascade; IEA.	
GO; GO:0006466; Protein amino acid phosphorylation; IEA.	
InterPro; IPR00719; Prot_kinase.	
InterPro; IPR00986; SH3.	
InterPro; IPR01455; SH3.	
InterPro; IPR01245; Tyr_pk kinase.	
InterPro; IPR008266; Tyr_pk kinase_AS.	
Pfam; PF00069; Pkinase_1.	
Pfam; PF00017; SH3; 1.	
Pfam; PF00018; SH3; 1.	
PRINTS; PR00401; SH3DOMAIN.	
PRINTS; PR4442; SH3DOMAIN.	
PRINTS; PR00109; TYRKINASE.	
ProDom; PD0000001; Prot_kinase_1.	
ProDom; PD0000093; SH2_1.	
ProDom; PD000066; SH3; 1.	
SMART; SN00252; SH3; 1.	
SMART; SN00326; SH3; 1.	
SMART; SN0215; TYRKC; 1.	
PROSITE; PS000107; PROTEIN KINASE ATP; 1.	
PROSITE; PS50001; PROTEIN KINASE DOM; 1.	
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.	
PROSITE; PS50001; SH2; 1.	
PROSITE; PS50002; SH3; 1.	
ATP-binding; Kinase; SH3 domain; Transferase; Tyrosine-protein kinase.	
NON_TER	509
SEQUENCE	509 AA; 58253 MW; 50BBC6407D109519 CR64;
Query Match	24.2%
Best Local Similarity	41.4%
Matches	67; Conservative 27; Mismatches 64; Indels 4; Gaps
38	VALGSPAGGPAGPAGPSLRGGPLTVSEDDWVTLSEVSGRENNIPVHGRVS --- HG
67	TALHYSPEPSHDGDLGFEKGEHLRLLEQNGEWKKAQSLTGGQGFVNPVAKANSLEPEP
94	WLYEGGLSREKAELLILLPGNPGAAFLIREQSQTGGSYSLSRSLRSRPAWDIIRTRHICL
127	WPFKONLSKRDQERQLLAPGATHFSFLINESSTAGSFPSLSPDQDQNQGVVKHYKIRNL
154	DNGWLYISPLRITFPSSLQALWDYHSELADDCCILKEPVCLQR 195
187	DNGGFIYSPITFSGLHELYRHYNASDGLCTRSLSPCQTQK 228

DE Proto-oncogene FYN.  
 GN P59YN.  
 OS Rattus norvegicus (Rat).  
 OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Brain.  
 RA Nemoto K., Sekimoto M., Kageyama H., Fukunachi K., Nemoto F.,  
 Ueyama T., Senba E., Tomita I.;  
 RT Expression of three type mRNAs encoding rat FYN by alternative  
 splicing within 5' untranslated region; Submitted within "5' untranslated region";  
 RL to the EMBL/Genbank/DBJ databases.  
 CC !- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; U35365; AAB2942.1; -.  
 DR PIR; PT0199; PT0199.  
 DR HSSP; P06241; 1FYN.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:004713; F:protein-Tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0007242; F:intracellular signaling cascade; IEA.  
 DR GO; GO:0006686; F:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPRO00719; Prot\_Kinase.  
 DR InterPro; IPRO00980; SH2.  
 DR InterPro; IPRO01450; SH3.  
 DR InterPro; IPRO01244; Tyr\_pk kinase.  
 DR InterPro; IPRO08266; Tyr\_pk kinase\_AS.  
 DR Pfam; PF00069; Pkinase\_1.  
 DR Pfam; PF00017; SH2\_1.  
 DR Pfam; PF00018; SH2\_1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00442; SH3DOMAIN.  
 DR PRINTS; PR00109; TURKINASE.  
 DR ProDom; PD000093; SH2\_1.  
 DR ProDom; PD000066; SH3\_1.  
 DR SMART; SN00252; SH2\_1.  
 DR SMART; SN00326; SH3\_1.  
 DR SMART; SN00219; TyrKc\_1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS50001; SH3\_1.  
 DR PROSITE; PS50002; SH3\_1.  
 KW ATP-binding; Kinase; SH3 domain; Transferase; Tyrosine-protein kinase.  
 SEQUENCE 537 AA; 60702 MW; 1IAE420919BPF1 CRC64;

Query Match 23.3%; Score 315.5; DB 11; Length 537;  
 Best Local Similarity 38.3%; Pred. No. 9e-20; Gaps 3;  
 Matches 77; Conservative 24; Mismatches 81; Indels 19; Gaps 3;

QY 10 SLPSPLSSSVQGQPYTMKAERSKATA-----VALGSPAGGPAAELSLRL 55  
 Db 46 SIPNTNFHAGGQGLTVFGNVNSSHGTLRGGTGYVALDYBEARTEDDLSPHK 105  
 Qy 56 GEPLTIV-SEDGDWNTVLSVSVSGREBYNIPVHVGYV---SHGMLYEGISLREKAEILL 110  
 Db 106 GEKFQIINSSGDDWNTVLSVSVSGREBYNIPVHVGYV---SHGMLYEGISLREKAEILL 165  
 Qy 111 PGNPQGAFLIRESQTRGSYSLSVRLSRPASWDRHYRHICLDNGWLYISPLTFPSLQ 170  
 Db 166 FGNPRTELRLRESETKGAVSLSIRDWDMDKGDIYKHYKRKLQDNGGTTTRAQFETLQ 225  
 Qy 171 ALVDHYSLEADDCCLKEPC 191  
 Db 226 QLVQHYSERAGLCCRUVVPC 246

RESULT 14  
 O93411 PRELIMINARY; PRT; 496 AA.  
 ID O93411  
 AC O93411;

DT 01-NOV-1998 (TREMBLrel. 08; Created)  
 DT 01-NOV-1998 (TREMBLrel. 08; Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25; Last annotation update)  
 DE Non-receptor protein tyrosine kinase Ialoo.  
 OS Xenopus laevis (African clawed frog).  
 OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Amura; Mesobatrachia; Pipidae; Pippidae.  
 OC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Weinstein D.C., Marden J., Carnevali F., Hemmati-Briantou A.;  
 RT "RGF-mediated mesoderm induction involves the Src-family kinase  
 Ialoo.";  
 RL Nature 0;0-0(1998).  
 CC !- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; AF081803; AAC31209.1; -.  
 DR HSSP; P08531; 1ADS.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:004713; F:protein-Tyrosine kinase activity; IEA.  
 DR GO; GO:004872; F:receptor activity; IEA.  
 DR GO; GO:0016740; F:transerase activity; IEA.  
 DR GO; GO:0007242; F:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPRO00719; Prot\_amino acid phosphorylation; IEA.  
 DR InterPro; IPRO00980; SH2.  
 DR InterPro; IPRO01452; Tyr\_Pkinase.  
 DR InterPro; IPRO01245; Tyr\_Pkinase.  
 DR InterPro; IPRO08266; Tyr\_Pkinase\_AS.  
 DR Pfam; PF00069; Pkinase\_1.  
 DR Pfam; PF00017; SH2\_1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR0103; TURKINASE.  
 DR ProDom; PD000093; SH2\_1.  
 DR ProDom; PD000056; SH3\_1.  
 DR SMART; SM00252; SH2\_1.  
 DR SMART; SM00326; SH3\_1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS50001; SH2\_1.  
 DR PROSITE; PS50002; SH3\_1.  
 KW ATP-binding; Kinase; Receptor; SH3 domain; Transferase;  
 KW Tyrosine-protein kinase.  
 SEQUENCE 496 AA; 562.5 MW; 96223A6F99689965 CRC64;

Query Match 23.2%; Score 313.5; DB 13; Length 496;  
 Best Local Similarity 36.4%; Pred. No. 1.2e-19; Gaps 6;  
 Matches 78; Conservative 34; Mismatches 75; Indels 27; Gaps 6;

QY 1 MGLSPSR-----RKSLPSPLSSSVQGQGP--YTM-EAERS-----KATAVALG 41  
 Db 1 NGCIKSDNTTGKSPGPPEOTQHVDPTSTVMTKPPRSSGPREEQBEVYLALY 60  
 Qy 42 SPPAGSPABSLRLGPBLTVYSEGDWWTLSVSESGREYNNIPSYVGKV---SHGWLYE 97  
 Db 61 DYDGVRHEGDLTFRKGDFHLLKKESEGEWMEACLISTGEGVPSNTVAFNSLESBWEYFK 120  
 Qy 98 GLSREKABELLILLPGLNPGAFLIRRSQTRGSYSLSVRLRPASTDRIHYRHICLDNGW 157  
 Db 121 GMSRKERAEQLSPVNSKGAFMIRDSETMGCPFLSVR---DSQSDTVKHYKIRLDDGG 176

Qy 158 LYISPRITFPSPLOALYDHYSELADDICCLKEPC 191  
 Db 177 FFISTRIPFPSPSLPELVRYHQGKVDGLCQLCTIPC 210

RESULT 15  
 Q99SPW1

Search completed: February 20, 2004, 09:32:11  
 Job time : 90 secs

ID Q99PWI PRELIMINARY; PRT; 541 AA.  
 AC "Rattus norvegicus (Rat)."  
 DT 01-JUN-2001 (TRIMBLrel. 17, Created)  
 DT 01-JUN-2001 (TRIMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TRIMBLrel. 25, Last annotation update)  
 DE Protein tyrosine kinase c-Yes (Fragment).  
 GN C-YES.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TAXID=10116;  
 RN [1]\_RN  
 RP SEQUENCE FROM N.A.  
 RA Shimizu M.;  
 RA "Rat mRNA for protooncogene c-Yes.";  
 RT Submitted (JAN-2000) to the EMBL/GenBank/DDJB databases.  
 RL -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC EMBL: AB037472; BAB21451.1; -.  
 DR EMBL; HSSP: P00523; 2.PH.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:000413; F:protein-Tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR000980; SH2.  
 DR InterPro: IPR01452; SH3.  
 DR InterPro: IPR01245; Tyr\_Pkinase.  
 DR InterPro: IPR008266; Tyr\_Pkinase\_AS.  
 PFam: PF00069; kinase; 1.  
 PFam: PF00017; SH2; 1.  
 DR PRINTS: PR00401; SH2DOMAIN.  
 DR PRINTS: PR00452; SH3DOMAIN.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProdDom: PD000001; Prod\_kinase; 1.  
 DR ProdDom: PD000093; SH2; 1.  
 DR ProdDom: PD000066; SH3; 1.  
 DR SMART: SM00326; SH2; 1.  
 DR SMART: SM003219; TyrKc1; 1.  
 DR PROSITE: PS000107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS55011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS000109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS55001; SH2; 1.  
 DR PROSITE: PS55002; SH3; 1.  
 KW ATP-binding; Kinase; SH3; domain; Transferase; Tyrosine-protein kinase.  
 FT NON\_TER 541  
 SEQUENCE 541 AA; 60587 MW; B3634PF8F2C030DA CRCE4;  
 SQ

Query Match 22.9%; Score 309.5; DB 11; Length 541;  
 Best Local Similarity 32.0%; Pred. No. 3.2e-19;  
 Matches 81; Conservative 41; Mismatches 92; Indexes 39; Gaps 7;

Qy 38 VALQSFPRGGPAHLRIGEPLTVSE-DGDNWVTVLSEVSGREVNIPSHVGRV---SH 92  
 Db 95 VALYDYEARTTEDLSFKGERFQLINNTEGDWNWNEARS-TATGNGYIIPSNYVAPADSIQE 154  
 Qy 93 GMLYEGLGLREKEAFLLLPGNPGSAFLRESQPTGYSYLPSVLRSRSPASWDRTR----H 147  
 Db 155 ENTFEGKMRKDARPLLLPGNQRQFPLTRESEETKGASLTSIR---DWDETRGDNYKH 209  
 Qy 148 YRTHCLDNGWLYISPRLTPEPSLQLWDHYSELADDICCLLKEPC-----VLRQAGP 198  
 Db 210 YKFRKLDDNGGYITRAQFDTLQKVLRHYTEHDGGLCHKLTTVCPTVKPTQGLAKDAME 269  
 Qy 199 LPGKDKDIPLPVYQR-----TPLNNKEDLSSLLFSPATGSESSLSEGLRES 244  
 Db 270 IPRESLRLLEVKLGCFCGEVWMGTWNGTITKA1KTKPSTMMPAAFLQEAQIMKRLRDK 329  
 Qy 245 L-SPPYISLNDAEV 256  
 Db 330 Tryptophanyl 342

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_B2n model

Run on: February 20, 2004, 09:34:08 ; Search time 3092 Seconds  
 (without alignments)  
 368,645 Million cell updates/sec

Title: US-09-939-853A-75  
 Perfect score: 1 MGSILSRKSLPSPSISSV.....RESLSFYISLNDEAVSLDDA 261

Scoring table: BLOSUM62  
 Xgapop 10.0 , Xgapext 0.5  
 Ygapop 10.0 , Ygapext 0.5  
 FGapop 6.0 , FGapext 7.0  
 Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
 -MODUL=frame+P2n.mode1 -DEV=xlp  
 -Q=Cgen2.1/USP0I spoql.p/USP0I.spoql.p/runat\_19022004\_145339\_24464/app/query.fasta\_1.455  
 -DB=Cgen2.1/USP0I spoql.p/USP0I.spoql.p/runat\_19022004\_145339\_24464/app/query.fasta\_1.455  
 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
 -DOCAALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODDB=LOCAL  
 -OUTFWMT=>to -NORM=ext -HEAPSIZE=500 -MINLEN=2000000000  
 -USER=US09-939-853@CGN.1 -MAXLEN=2000000000  
 -NO MMAP -LARGEQUERY -NEGSCORE=0 -WAIT -DSBLOCK=100 -LONGLOG  
 -DEV TIMEOUT=30 -WARN THREADES=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
 -FGAEEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DROPOF=6 -DELEXT=7

Database : GenEmbl:

Result No.	Query	Score	Match	Length	DB ID	Description
1	em_vi:*	1353	100.0	1183	6 AX44133	AX44133 Sequence
2	em_htg_hum:*	1353	100.0	1183	6 AX44133	AX44133 Sequence
3	em_htg_other:*	1347	99.6	786	6 AX511153	AX511153 Sequence
4	em_htg_mus:*	1347	99.6	786	6 AX512845	AX512845 Sequence
5	em_htg_fod:*	1347	99.6	786	6 AF290985	AF290985 Homo_sapi
6	em_htg_mam:*	1347	99.6	786	9 AF326353	AF326353 Homo_sapi
7	em_sy:*	1347	99.6	786	9 AK05645	AK05645 Homo_sapi
8	em_htg_musi:*	1347	99.6	2538	9 BC02041	BC02041 Homo_sapi
9	em_htg_other:*	1347	99.6	2567	6 AX42880	AX42880 Sequence
10	em_sy:*	1347	99.1	2788	6 AX780857	AX780857 Sequence
11	em_vi:*	1196.5	88.4	737	6 AX511155	AX511155 Sequence
12	em_htg_hum:*	1196.5	88.4	737	9 AF290986	AF290986 Homo_sapi
13	em_htg_other:*	1196.5	88.4	737	9 AX511151	AX511151 Sequence
14	em_htg_mus:*	1028	76.0	1348	6 AX511150	AX511150 Sequence
15	em_htg_fod:*	1028	76.0	1348	10 BC052655	BC052655 Mus muscu
16	em_htg_mam:*	1028	76.0	1384	10 AF344990	AF344990 Mus muscu
17	em_sy:*	1019	75.3	1321	10 BC056035	BC056035 Xenopus laevis
18	em_htg_musi:*	583	41.1	1571	5 UY217759	UY217759 Rattus norvegicus
19	em_htg_other:*	501	37.0	1631	10 MMU29056	MMU29056 Mus musculus
20	em_sy:*	492	36.4	921	10 BC03222	BC03222 Mus musculus
21	em_vi:*	492	36.4	2648	10 AJ31377	AJ31377 Mus musculus
22	em_htg_hum:*	492	36.4	2773	10 MMU13177	MMU13177 Mus musculus
23	em_htg_other:*	491.5	36.3	841	10 AY079449	AY079449 Mus musculus
24	em_sy:*	491.5	36.3	841	10 AY079450	AY079450 Mus musculus
25	em_vi:*	487	36.0	1076	9 HSU30473	HSU30473 Homo sapien
26	em_htg_hum:*	487	36.0	1870	9 BC007042	BC007042 Homo sapien
27	em_htg_other:*	487	36.0	2021	9 HSU44403	HSU44403 Human_src-1
28	em_htg_mus:*	487	36.0	2109	6 AX428893	AX428893 Sequence
29	em_htg_fod:*	487	36.0	2665	6 AX3333017	AX3333017 Sequence
30	em_sy:*	487	36.0	2665	9 D89077	D89077 Homo sapien
31	em_vi:*	487	36.0	3090	9 HSM80530	HSM80530 Human DNA
32	em_htg_hum:*	482	35.6	145833	2 AC065339	AC065339 Human DNA
33	em_htg_other:*	482	35.6	145833	5 AY278230	AY278230 Gallus gallus
34	em_htg_mus:*	467.5	34.6	849	5 M57697	M57697 Mouse lynx norvegicus
35	em_htg_fod:*	380.5	28.1	2703	10 RATLYNBTYR	RATLYNBTYR Synthetic
36	em_sy:*	375.5	27.8	1538	10 RNAF000301	RNAF000301 Synthetic
37	em_vi:*	375.5	27.8	2302	10 RNAF000302	RNAF000302 Synthetic
38	em_htg_hum:*	375.5	27.8	2824	10 AY335782	AY335782 Synthetic
39	em_htg_other:*	370.5	27.4	1539	12 AR380613	AR380613 Sequence
40	em_htg_mus:*	370.5	27.4	2298	6 AX779985	AX779985 Sequence
41	em_sy:*	370.5	27.4	2298	6 M1638	M1638 Human lynx norvegicus
42	em_vi:*	370.5	27.4	2298	6 AX779984	AX779984 Sequence
43	em_htg_hum:*	370.5	27.4	4175	6 Li14782	Li14782 Rattus norvegicus
44	em_htg_other:*	370	27.3	1601	10 RATLYNATYR	RATLYNATYR Synthetic
45	em_sy:*	370	27.3	2041	10 RNAF000300	RNAF000300 Synthetic

## SUMMARIES

## DESCRIPTION

RESULT 1  
 ALIGNMENTS



















Qy	81 AsnIleProSerValHisValGlyIleSerHisGlyTrpLeuIleGlyLeuSer 100	Oy	1 MetGlySerLeuProSerArgArgLysSerIleuProSerLeuSerSerIleu 20
Qy	238 CACATGCCAGTGTTGATGGCTAAAGTCGGCACGGTGGCTGAGGGCTGAGC 297	Db	282 ATGGAGATTTGTCACAGGGAAACCCAGGGCTTCCTCATC 338
Db	101 ArgGluIysAlaGluGluLeuIleLeuIleProGlyIysAlaPheLeuIle 120	Oy	21 GlyGlyIysGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaLeu 40
Db	298 CGGAGAAGCGGAGACTTCCCTTAATGGGAAACCCGGAGGGCTTCCTCATC 357	Db	339 CCAGACAGAACCCGGTGTCCATTGAAAGAACAGAAAGTCAGTGAGGCCTG 398
Qy	121 ArgGluIysSerGlySerIysSerLeuSerValArgLeuSerArgProAla 140	Oy	41 GlySerIysProAlaGlyGlyProAlaGluIleSerIleuArgLeuGluProLeuThr 60
Db	358 CGGAGAACGCCCCACCGAGCTGCTTATCCCTGCGACTCGCGCCCTGCA 417	Db	399 GGCGATTCAGCTCCAGGAAACGGCCAGACTATCTGAGACTGGAGGGCTGAC 458
Qy	141 SerTPAsPArgIleArgHistYraGlyLeuIleScylsCysLeuIlePAsnGlyTrpLeuIle 160	Oy	61 IleValSerGluAspGlyAspPTPThrValLeuSerGluValSerGlyArgGlyIys 80
Db	418 TCTTGGGACCCGATCAGACAATCAGGATAACAGGATAACGGCTTGTGAAATGCTGGTGTACATC 477	Db	459 ATCATCTGAGGATGGAGATGGGAACTGCTGGAACTGCTGGAGAGTGTAC 518
Qy	161 SerProArgLeuThrProSerLeuIleGlnAlaLeuValAspIstYrsSerGluLeuAla 180	Oy	81 AsnIleProSerValHisValGlyIysValSerHisGlyYtPheLeuIleGlyLeuSer 100
Db	478 TCACTCGCCACACTTCCCCTACTCCACGGCTGGAGCATACTCTAGTAGCA 537	Db	519 CACATGGCCAGTGTGTGTGCTAAAGTGCCTGAGGAGGGCTCTCTCATC 578
Qy	181 AspAspIleCysCysLeuLeuIleGlyLeuIleProLeuPro 200	Oy	101 ArgGluIysAlaGluGluLeuIleLeuIleProGlyIysAlaPheLeuIle 120
Db	538 GATGGCATCTGCTGCCCCCTAGGGAGCCGGTGTGCTGAAAGCTGGGCCACTACCT 597	Db	579 CEGGAGAAAGCCGGAGGACTACTCTGTACCTGGAAACCCGGAGGGCTCTCATC 638
Qy	201 GlyIysAspIleProLeuProValThrValGlnArgThrProLeuAsnTriPheGluIle 220	Oy	121 ArgGluIserGlnThrArgArgGlySerIysSerLeuIleSerArgLeuSerArgProlAla 140
Db	598 GGCAGAAAGATAACCTTCCACCTGTGACTGTGCCAACATCATACTAAATTGGAAAAGCTG 657	Db	639 CGGGAGGGCCAGACCCAGAGGGAGGCTGGCTTACCTGTCGTCGACTAGGCCCTGGAA 698
Qy	221 AspSerSerLeuLeuPhasIerGluAla--AlaThrGlyGluIleSerGlu 239	Oy	141 SerTPAsPArgIleAspHistYtaArgIleHisCysIleuAspAsnGlyTrpLeuIle 160
Db	658 GACCGCGAGCTCTGTTGAGGACCTGGAGCTGAGTGGGGAGCATCTCTGTCAGTGAG 717	Db	639 TCTGGGACCGGATTCAGACACTACAGGATAACGGCTTGTGACATGGCTGCTACATC 758
Qy	240 GlyLeuIargGluIleSerIerYtIleSerLeuAsnAspIleAlaValIleSerIleuAsp 259	Oy	161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspIstYrsSerGluLeuAla 180
Db	718 GGGCTCCGAGTCAGTCCCTAGTCATGCTGAGGAC----CCCTTGAT 771	Db	759 TCACTCGCCCTACCTTCCACCTTCCACCTCCACGCTTGTGGAGCTTGTGCTGAGGCTTGTGCTAC 818
Qy	260 AspAla 261	Oy	181 AspAspIleCysIleuLeuIleProLeuAsnTriPheGluIle 200
Db	772 GATGCT 777	Db	819 GATGGCATCTGCTGTCCTCAGGGAGCTGGCTGCTGAGGCTTGTGGCCACTACTACCT 878
Qy	RESULT 14 AX511150 AX511150 1348 bp DNA linear PAT 27-SEP-2002	Oy	201 GlyIysAspIleProLeuProValThrValGlnArgThrProLeuAsnTriPheGluIle 220
DEFINITION	Sequence 1 from Patent WO0242452.	Db	879 GGAAAAAGATAACCTCCTGACTGACTGGAAACATATCACTAAATGGAAAGCTG 938
ACCESSION	AX511150	Qy	221 AspSerSerLeuLeuPhasIerGluAla--AlaThrGlyGluIleSerIleuSerGlu 239
VERSION	AX511150.1 GI:23392044	Db	939 GACCGAGCCCTCTGTTCTGGAGAACCTCCAGTGGGAGGCTCATCTGCTAGTGAG 998
SOURCE	Mus musculus (house mouse)	Qy	240 GlyLeuIargGluIleSerIerYtIleSerLeuAsnAspIleAlaValIleSerIleuAsp 259
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Db	999 GGCTCTGGAGAGTCCTCAGTTCCTGAGCTGGCTGAGGAC-----CCCTTGAT 1052
REFERENCE	1. McGlade, J.C. and Loreto, M.P.	Oy	260 AspAla 261
AUTHORS	McGlade, J.C. and Loreto, M.P.	Db	1053 GATGCT 1058
JOURNAL	Patent: WO 0242452-A 1 30-MAY-2002; The Hospital for Sick Children (CA)	RESULT 15 AF2B7467	RESULT 15 AF2B7467
FEATURES	Location/Qualifiers 1. .1348 /organism="Mus musculus" /mol_type="unassigned DNA" /db_xref="taxon:10090"	LOCUS AF2B7467	LOCUS AF2B7467
source		DEFINITION Mus musculus (house mouse)	DEFINITION Mus musculus Sro-like adaptor protein-2 mRNA, complete cds.
		VERSION AF2B7467.1 GI:17351918	VERSION AF2B7467.1 GI:17351918
		KEYWORDS	KEYWORDS
		ORGANISM	ORGANISM
		Mus musculus	Mus musculus
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
		REFERENCE 1. Loreto, M.P., Berry, D.M. and McGlade, C.J.	REFERENCE 1 (bases 1 to 1348) Loreto, M.P., Berry, D.M. and McGlade, C.J.
		AUTHORS	AUTHORS
		TITLE Functional cooperation between c-Cbl and Src-like adaptor protein 2 in the negative regulation of T-cell receptor signaling	TITLE Functional cooperation between c-Cbl and Src-like adaptor protein 2 in the negative regulation of T-cell receptor signaling
		DB: JOURNAL Mol. Cell. Biol. 22 (12), 4241-4255 (2002)	DB: JOURNAL Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
		ORIGIN	ORIGIN
		Alignment Scores:	Alignment Scores:
Pred. No.:	1.51e-75	Length:	1348
Score:	1058.00	Matches:	208
Percent Similarity:	85.50%	Conservative:	16
Best Local Similarity:	79.39%	Mismatches:	34
Query Match:	75.98%	Indels:	4
DB:	6	Gaps:	3

PUBMED 12024016  
 REFERENCE 2 (bases 1 to 1348)  
 AUTHORS Loreto, M. P. and McGlade, C. J.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada  
 FEATURES Location/Qualifiers  
 source 1 .1348  
 /organism="Mus musculus"  
 /mol\_type="RNA"  
 /db\_xref="taxon:10090"  
 /dav\_stage="day 15 embryo"  
 282 .1061  
 /note="SLAP-2"  
 /codon\_start\_1  
 /product\_id="AAI18196\_1"  
 /protein\_id="GI:17351919"  
 /translation:"MGLISRSRGPPLTISEDGMWWTQSEVREYIMPSVUYAKTAHGNLYEGRK  
 AELILLGPNPAGFLRESQRPGCTLSVRLRSPLASWDEIRHTRIQRLDNGWYIS  
 PRITFSPHSHAEVHSELAQGICCPURPCPVQKLQPLPGKDTPPPVTVPTSVSNWKK  
 LRSLSLIFEARSGEASLSEGIRESSYSSYSLAEDPLDDA"  
 misc\_feature 381 .551  
 /note="Region: SH3 domain"  
 misc\_feature 552 .824  
 /note="Region: SH2 domain"  
 misc\_feature 1301 .1306  
 ORIGIN  
 polyA\_signal

---

Alignment Scores:  
 Pred. No.: 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Score: 1028.00 Length: 1348  
 Percent Similarity: 85.50% Matches: 208  
 Best Local Similarity: 79.39% Conservative: 16  
 Query Match: 75.98% Mismatches: 34  
 DB: 10 Indels: 4  
 Gaps: 3

US-09-939-853A-75 (1-261) x AF287467 (1-1348)

Qy 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 282 ATGGGAGTTGTCCAGCAAGGAANAC...-TCCAGGCCACCCAGCCTCTCGT 338  
 Qy 21 GlngGlyGlnGlyProValThrMetGluAlaGluGlySerLysAlaLeu 40  
 Db 339 CCAGACAGGAACCCGTTCATGAACAGAACAGAACAGTCAGCTGAGCCCTG 398  
 Qy 41 GlySerPheProAlaGlyGlyProAlaGlyLeuSerLeuGlyLeuProLeuThr 60  
 Db 399 GGCAAGTTCCAGAGGTGACAGGCCAAACTATCTCAAGTCAGTCGGGGCTGAC 458  
 Qy 61 IleValSerGluGlyGlyAsparticProlineSerGluValSerGlyGluThr 80  
 Db 459 ATCACCTCTGAGATGGATGGAAATTGCGGAAGTCAGTCAGTCAGCAAGGAGAAC 518  
 Qy 81 AsnLeProSerValHisValGlyLysValSerHisGlyLeuSerGlyGluThr 100  
 Db 519 CACATGCCAGCTGTATGCTAAAGGCCCAAGGGCTGTACAGGGCCCTGAGC 578  
 Qy 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAlaAlaLeuLeu 120  
 Db 579 CGGGAAAGCGAAACTCTCGTACCTGGAAACCCGAGGGCTTCCTCATC 638  
 Qy 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 639 CGGGAAAGCCGACCGAGGAGGAGGCTGATCCGGTCCCTCCACTCGCCGCCCTGCA 698  
 Qy 141 SerTPAspArgIleArgHistidineArgLysIleAspArgLysLysLeuLeu 160  
 Db 699 TCTTGGGACCCGATCAAGCAACTACAGATCACGGCTCTGCTACAATGGCTGGCTACATC 758  
 Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:33:07 ; Search time 368 Seconds  
(without alignments)

3012.988 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLPSRKSLPPSPLSSV.....RESLSFYISLNDBAVSLDDA 261

Scoring table: BLASTN62

xgapop 10.0 Xgapext 0.5

ygapop 10.0 Ygapext 0.5

fgapop 6.0 Fgapext 7.0

Delop 6.0 Deletx 7.0

Searched: 3373863 seqs., 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+P2n.model -DEV=xplp
-Q=Cgn2/1/USP1I.spo1_P2n.rng -OFILE=fastDP -SUFFIX_P2n.rng -MINMATCH=0.1 -LICOPOL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCAIGN=200 -THR SCORE=PC -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -THEAPSITE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=USC9939853@CCGN_1 -1708@runt_19022004_145338_24420/app_query.fasta_1.455
-NOBLOCK=1/USP1I.spo1_P2n.rng -MINMATCH=0.1 -LICOPOL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCAIGN=200 -THR SCORE=PC -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -THEAPSITE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=USC9939853@CCGN_1 -1708@runt_19022004_145338_24420 -NCPU=6 -ICPU=3
-NOBLOCK=1/USP1I.spo1_P2n.rng -MINMATCH=0.1 -LICOPOL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCAIGN=200 -THR SCORE=PC -THR MAX=100 -THR MIN=0 -ALIGN=15
-USER=USC9939853@CCGN_1 -1708@runt_19022004_145338_24420 -NCPU=6 -ICPU=3
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -TREADS=1 -XGAPPOP=1.0 -XGAPEXT=0.5 -FGAPOP=6
-FOGAPEXT=7 -YGAPPOP=1.0 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

## RESULT 1

ABK61465

standard; cDNA; 1183 BP.

ID ABK61465

XX

AC ABK61465;

XX DT 18-JUN-2002 (first entry)

XX DE Human cDNA encoding protein NOV13.

XX Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;

KW cell signal processing disorder; metabolic pathway modulation disorder;  
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate, cancer;  
KW uterus; cancer; immune response; graft-versus-host disease;  
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;  
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;  
KW Albright hereditary osteodystrophy.  
KW OS Homo sapiens.

## SUMMARIES

XX PN WO200216539-A2.

XX PD 28-FEB-2002.

XX PP 27-AUG-2001; 2001WO-US036510.

XX PR 25-AUG-2000; 2000US0228191P.

PR 08-FEB-2001; 2001US0267300P.

PR 20-FEB-2001; 2001US026961P.

PR 20-MAR-2001; 2001US0277337P.

XX PA (CURA-) CURAGEN CORP.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Database : N Geneseq\_29Jan04:\*

1: geneseqn1980s:\*

2: geneseqn1990s:\*

3: geneseqn2000s:\*

4: geneseqn2001s:\*

5: geneseqn2001bs:\*

6: geneseqn2002s:\*

7: geneseqn2003s:\*

8: geneseqn2003bs:\*

9: geneseqn2003bs:\*

10: geneseqn2004s:\*

\* SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1353	100.0	1183	6 ABK61465	
2	1347	99.6	786	6 AAI44089	Aai44087 Human ORF
3	1347	99.6	786	6 ABQ99151	Abq99151 Human ORF
4	1269	99.6	2567	6 ADA43980	Ada43980 Human Src
5	1269	93.8	837	3 AAC77202	AAC77202 Human ORF
6	1206.5	89.2	1413	6 ABQ99374	Abq99374 Human cod
7	1196.5	89.4	737	6 AAI44090	Aai44090 Mouse M2R
8	1128.5	83.4	2049	5 AAS74750	Aas74750 DNA encod

QY	121	ArgGlyserGlnThrArgArgGlyserSerSerLeuSerValArgLeuSerArgProAla	140
Db	758	CGGGAAAGCCGACGGAGGGCTTACTCTCTCAGTCGGCTTCAGCGGCCGCA	817
QY	141	SerIlePAspIleArgHistYArgIleHisCysLeuAspAsnGlyTrpLeuTyrrile	160
Db	818	TCTCTGGGACCCGATCGAGACATACAGGATCCACTGCTGCCTGTCACATC	877
QY	161	SerProArgLeuThrProSerLeuGlnAlaLeuValAspHistYArgSerIleUrea	180
Db	878	TCAACGGGGCTCACCTTCCCCTCACTCCAGCCCCGTTGACCATTA	937
QY	181	AspAspIleCysCysteIleLeuProValThrValGlnArgIleGlyProLeuPro	200
Db	938	GATGACATCTGTGCTACTAGGGCCCTGTGCTGAGGGCTGCCGCCCTCC	997
QY	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
Db	998	GCGAAGATATACTACCCCTACCTGTGACTGTGAGAGGACACACTGCAAGGCTG	1057
QY	221	AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerIleLeuGluGly	240
Db	1058	GACAGCTCCCTCTCTGTTCTAAAGCTGCCAGGGAGGGTCTCTAGTGAGGT	1117
QY	241	IleArgGluSerLeuSerPheTyrIleSerLeuAspGluAlaValSerLeuAspAsp	260
Db	1118	CTTCCGGGAGTCCCTCAAGCTTCTAACATCAAGCCTGAATGACGGCTCTTGGATGAT	1177
QY	261	Ala	261
Db	1178	GCC	1180
RESULT 2			
	AAI44089		
ID	AAI44089	standard; cDNA;	786 BP.
XX			
AC	AAI44089;		
XX			
DT	03-OCT-2002	(first entry)	
XX			
DE	Human modulator of antigen receptor signalling protein coding sequence.		
XX			
KW	Human; gene; ss; gene therapy; modulator of antigen receptor signalling;		
KW	MARS; tumour suppressor gene; Scrl-like adaptor protein; SLAP;		
KW	myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;		
KW	immunosuppression; myeloproliferative disorder; breast cancer.		
XX			
XX	Homo sapiens.		
OS			
FH	Location/Qualifiers		
XX			
CD5	1.	.786	
FT	/*tag= ^ /product= "Human MARS protein"		
XX			
PN	WO200242452-A2.		
XX			
PD	30-MAY-2002.		
XX			
PT	26-NOV-2001;	2001WO-CR001662.	
XX			
PR	27-NOV-2000;	2000CA-02324663.	
XX			
PA	(HOOP-) HOSPITAL FOR SICK CHILDREN		
XX			
PI	McGlade JC,	Loreto MP;	
XX			
DR	WPI:	2002-566564/60.	
P-PDB:	AA013457.		
XX			
PT	New isolated modulator of antigen receptor signalling protein or its fragment, useful for treating malignant disorders such as myeloid malignancies, autoimmune disorders and myeloproliferative disorders.		
PT			



(II) is useful in the treatment of immunodeficiency disorders, such as acquired immunodeficiency syndrome (AIDS), for the prevention and treatment of acute inflammatory disorders, chronic inflammatory disorders, autoimmune disorder and transplant rejection.

Sequence 786 BP; 162 A; 234 C; 159 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No. :	4.79e-109	Length:	786
Score:	134.00	Matches:	260
Percent Similarity:	99.12%	Conservative:	0
Best Local Similarity:	99.52%	Mismatches:	1
Query Match:	99.56%	Indels:	0
DB:	6	Gaps:	0
US-09-939-853a-75 (1-261) x ABQ74343 (1-786)			
Qy	1	MetGlySerIeuProSerArgGlySerLeuProSerProSerLeuSerSerSerVal	20
Db	1	ATGGCAGTGTGCCAGCAAGAAAGAAATTCTGCAACCCAAGCTTGAGTCCTCTGTC	60
Qy	21	GlnGlyGlnGlyProValThrMerGluAlaGluArgSerLysAlaAlaLeu	40
Db	61	CAAGGCAGGCCACCTGTGACCATGAGCATGAGAAAGAGAAACAGGCCAGGCCCTG	120
Qy	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerIeuArgLeuGlyGluProLeuThr	60
Db	121	GGCAGTTCCGGCAGGTGCCGCCGAGCTGCGCTGAGACTCGGGAGCCATTGACC	180
Qy	61	IleValSerGluAspGlyAspPTPPTPThrValLeuSerGluValSerGlyArgGluTy	80
Db	181	ATCGCTCTGAGATGGAGCTGGAGCTGCTGCTGAGCTCAGGCAAGCTCAGGAGCTAT	240
Qy	81	AsnIleProSerValHisValGlyValSerIleGlyTrpLeuThrGlyLeuSer	100
Db	241	AACATCCCAGGTCCAGGGCCAAAGCTCCATGGTGGCTGTGATGAGGCCCTGAGC	300
Qy	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlapheLeuIle	120
Db	301	AGGGAGAACGAGGAACCTGCTGTAACTGGAAACCTCTGGAGGGCCTCCCTCATC	360
Qy	121	ArgGluSerGinThrArgArgGlySerIeuSerValArgLeuSerArgProAla	140
Db	361	CGGGAAAGCG2AGCACGGAGAACAGGAACTGGCTTAACTCTGTCAGTCGGCCCTCTGCA	420
Qy	141	SerTrpAspArgIleArgHistYArgIleHistYsLeuAspAsnGlyTriPheTyrIle	160
Db	421	TCTGGGACCGATCTGACTCTGGATCACCTGGCTTACATGGCTGCTGATCATC	480
Qy	161	SerProArgLeuIlePheProSerLeuGlnAlaLeuValAspHistYSerGluLeuAla	180
Db	481	TCACCGCGCTCACCTTCCCTACTCCAGCCCTGTCGTCACCTACTGAGCTGCCG	540
Qy	181	AspAspIleCysLeuIleLysGluProCysValLeuGlnAlaGlyProLeuPro	200
Db	541	GATGACATCTGCTGCTACTCAAAGAGCCCTGTCAGGGCTGGCCGTCCT	600
Qy	201	GlyLYAspIleProLeuProValThrValGlnArgIleProLeuAsnTrpYsGluIeu	220
Db	601	GGCARGGATATACTCTACTCTACTGTCGTCAGGAGACTCAACTGGAAAGAGCTG	660
Qy	221	AspSerSerIeuLeuPheSerGluAlaIleIleSerLeuAspIeuAspIeu	240
Db	661	GACAGCTCCCTCTGTTCTGTCAGCTGCACTGTCAGCTGAGCTGAGCTGAGGT	720
Qy	241	IeuArgLysLeuSerLeuSerPheTyrIleSerLeuIeuAspGluAspIeuAspIeu	260
Db	721	CJCCGGAGTCCTAGCTCATCATGAGCTGATGACCTGAACTCATGAGCTTGGATGAT	780
Qy	261	Ala 261	
Db	781	GCC 783	

## RESULT 4

AAD41980  
ID AAD41980 standard; CDNA; 2567 BP.

XX  
XX

AC  
AC

D1  
D1

13 -DBC-2002 (First entry)

XX Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.

DB XX Human; SH2 /SH3-domain-containing adapter; Src-Like Adapter Protein-2; Score: 134.00

KW Human; SH2 /SH3-domain-containing adapter; signal transduction; autoimmune disease; cancer; immune disorder; signal transduction; rheumatoid arthritis; psoriasis; inflammation disorder; allergy; Crohn's disease; systemic lupus erythematosus; inflammatory bowel disease; gene therapy; multiple sclerosis; asthma; acute respiratory distress syndrome; dermatological disorder; pulmonary disorder; organ rejection; neuroprotective gene; ss.

XX Human SLAP-2

XX Human SLAP-

Alignment Scores:							
Aligned Pairs:	Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:		2.19e-108	1347.00	99.62%	99.62%	0	0
Percent Similarity:							
Best Local Similarity:							
Query Match:							
DB:							
US-09-939-853A-75 (1-261) x AAD43980 (1-261)		2567	260	0	0	0	0
QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal 20	XX	XX	XX	XX	XX
Ddb	415	ATCGGAACCTGCCAGAGAAATACTCTGCAGCCAAAGCTTGACTTCCTCTGTC 474	XX	XX	XX	XX	XX
Qy	21	GlnGlyGlnGlyProValThrMerGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40	XX	XX	XX	XX	XX
Ddb	475	CAGGCCGGGGACTGTGACCATGARGCAGGAGCAAGGCCACAGCGTGCCTCTG 534	XX	XX	XX	XX	XX
Qy	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuLeuArgLeuGlyGluProLeuThr 60	XX	XX	XX	XX	XX
Ddb	535	GGAGATTCCTCCGCAGGGCCGGCGAGCTGCGTGAGCTGGGGAGCCATGACC 594	XX	XX	XX	XX	XX
Qy	61	IleValSerGluaspGlyAspIrrPrpThrValLeuSerGluValIserGlyArgGlyIrr 80	XX	XX	XX	XX	XX
Ddb	595	ATGGCTCTGAGATGAGACTGGAGCTGGAGCTGGTGACGGGCTGCTGAGTCAGGGAGAT 654	XX	XX	XX	XX	XX
Qy	81	AspIleProSerValHisValGlyLysValSerHisGlyLysTrpLeuIrrGluGlyIleuSer 100	XX	XX	XX	XX	XX
Ddb	655	AACTATCCCAGCTCCAGTCAGCTGCCAAAGTCCTCCATGGGCTGTTCTCATC 714	XX	XX	XX	XX	XX
Qy	101	ArgGluIlysAlaGluGluLeuLeuLeuProGluAsnProGlu1GlyAlaIpheIle 120	XX	XX	XX	XX	XX
Ddb	715	AGSGAGAAGGAGAGGAACCTGGAACTGGTGTCTACCTGGGACCCCTGGCTTCATC 774	XX	XX	XX	XX	XX
Qy	121	ArgGluSerGlnThrArgGlySerIrrSerIrrSerLeuSerArgProAla 140	XX	XX	XX	XX	XX
Ddb	775	CGGGAGCCAGACAGAGGCTCTACTCTCTGAGCTGGCTCTGGCTGGCTGGCA 834	XX	XX	XX	XX	XX
Qy	141	SerIrrPaspArgLleIrrGlySerIrrSerIrrSerLeuSerArgLeuSerArg 160	XX	XX	XX	XX	XX
Ddb	835	TCCCGGGACGGATCAGACATACAGGATCCACTGCTGGCTGACATGGCTGACATC 894	XX	XX	XX	XX	XX
Qy	161	SerIrrArgLeuThrPheProSerLeuIrrAlaLeuValAlaLeuValAlaLeuAla 180	XX	XX	XX	XX	XX
Ddb	895	TCACGGCGCTCACTCTCCCTACCTCGGCTCTGCTGACATTACTCTGAGCTGGCG 954	XX	XX	XX	XX	XX
Qy	181	AspAspIleCysCysLeuLeuIrrGluProCysValLeuIrrArgAlaIrrGlyProLeuPro 200	XX	XX	XX	XX	XX
Ddb	955	GATGACATCTGCTCCACTGACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1014	XX	XX	XX	XX	XX
Qy	201	GlyIysAspIleProLeuProValThrIrrValGlnIrrArgThrProLeuAsnTrpLysGluLeu 220	XX	XX	XX	XX	XX
Ddb	1015	GCGAAGGATAATCCCTACTGTGACTGAGGACACCACTAACATGAAAGAGCTG 1074	XX	XX	XX	XX	XX
Qy	221	AspSerIrrSerLeuSerGluAlaIrrIrrGlyGluGluSerGluLeuSerGluGly 240	XX	XX	XX	XX	XX
Ddb	1075	GACGCTCTCTCCCTGTTCTGAAAGCTGCCACAGGGAGACTCTCTGAGTGGGT 1134	XX	XX	XX	XX	XX
Qy	241	LeuIrrGluSerLeuSerPhenylIrrSerLeuAsnAspGluAlaValSerLeuAspAsp 260	XX	XX	XX	XX	XX
Ddb	1135	CTCCGGAGCTCCCTAGCTTAATGAGCTGAAATGAGCTGAAATGAGCTGAAATGAGCTG 1194	XX	XX	XX	XX	XX
Qy	261	Ala 261	XX	XX	XX	XX	XX
Ddb	1195	GCC 1197	XX	XX	XX	XX	XX
RESULT 5							
AAC77202	AAC77202 standard	cDNA;	837 BP.				
AAC77202	AAC77202						

Query Match:	93.79%	Indels:	0	Gaps:	0			
DB:	3							
US-09-939-853a-75 (1-261) × AAC77202 (1-837)								
Qy	15 SerIleSerSerSerValGlnGlyGlnGlyProValThrMetGluAlaGluArgSerLys	34						
Db	4 AGCTGTAGTTCCTCTGTCGAAGGCCAGGAACCTGAGCATGGAAAGGAGAAAG	63						
Qy	35 AlaThrAlaLeuGlySerIleProAlaGlyProAlaGlyLeuSerLeuLeuArg	54						
Db	64 GCCAGAGCGCCGCTGCGCAGTTCCGGCAGGTGCCCCGAGCTGCGTGCAGA	123						
Qy	55 LeuGlyGluProLeuThrIleAlaSerGluAspDlyAspTrpPhePheValLeuSerGlu	74						
Db	124 CTCGGGGCCATGGCCATGACCACTGAGTGGAGCTGGAGCTGGAGCTGTCAGA	183						
Qy	75 ValSerGlyArgGlutYrransIleProSerValAlaIleSerHisGlyTyrP	94						
Db	184 GTCTAGGGAGAGATAACATCCACGGTCAACGTTCAAGCTCCATGGTGG	243						
Qy	95 LeuTyGluGlyIleSerIleArgGlyLysAlaGluGluLeuLeuLeuLeuProGlyAspPro	114						
Db	244 CTGTATGGGCCCTAGCAAGGAGAAAGCTGCTGTTACCTTGAACTCT	303						
Qy	115 GlyGlyAlaIleLeuLeuArgGlySerIleSerIleSerLeuSerVal	134						
Db	304 GGAGGGCCCTCCCTATCGGGAGGCCAGGGCTCTATTCTGTGTCAGTC	363						
Qy	135 ArgLeuSerArgProAlaSerGlyAspAlaPheIleLeuGlySerLeuAsp	154						
Db	364 CGCCCTAGGGCCCTGCACTCTGAGACATAACAGGATCACAGCTGCTGAC	423						
Qy	155 AsnGlyTrpLeuItyrIleSerProArgLeuThrPheProSerLeuGlnAlaIleLeuAsp	174						
Db	424 AATGGCTGGGTGTCATCTACCGGCCTOACCTTCCCTACTCCAGGCCCTGGGAC	483						
Qy	175 HistYrSerGluLeuAlaAspDlyCysCysLeuLeuLysGluProCysValLeuGln	194						
Db	484 CATTACTCTAGCTGGGTGACATCTGCTACTAACAGGCTCTGTTCTGTCAG	543						
Qy	195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPro	214						
Db	544 AGGGCGGCCGCGCCAGATAAACCTTACCCCTGACTGTCGAGAGACACCA	603						
Qy	215 LeuAsnTrpIysGluLeuAspSerSerLeuLepheSerGluAlaIleArgIleGluGlu	234						
Db	604 CTCAACTGGAAAGCTGGCAGCTCCCTCTGAGTGCCTGAGTTCGAGGAGGAG	663						
Qy	235 SerIleLeuSerGluGlyIleArgGluSerLeuSerPheTyrylIleSerLeuLeuAspGlu	254						
Db	664 TCTCTCTCTAGCTGAGGTCTCCAGCTTACATCAGCTGATGACAG	723						
Qy	255 AlavalSerLeuAspAspAla	261						
Db	724 GCTGTCCTCTGGATGCC	744						
RESULT 6								
ABQ99374								
ID ABQ99374 standard; CDNA: 1413 BP.								
XX								
AC ABQ99374;								
XX DT 25-FEB-2003 (first entry)								
XX DE Human coding sequence SEQ ID 107.								
XX KW Human; expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human coding sequence SEQ ID 107.								
XX KW expressed sequence tag; EST; chromosome 20;								
KW haemopoietic disorder; central nervous system disease; viral infection;								
KW peripheral nervous system disease; non-healing wound; infectious disease;								
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;								
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;								
KW antiinflamatory; antihistaminic; immunosuppressive; neuroprotective;								
XX DE Human								







XX	27-JUN-2002.	Db	586	ACGGAGAAAGCAGGAACTGCCTGTGTTACCTGGAAACCCCTCGGGGCCCTCCATC	645
PD		Qy	121	ArgGluserGlnThrArgArgGlySerTyrSerLeuSerSerValArgLeuSerArgProAla	140
XX	30-MAY-2001; 2001US-00867550.	Db	646	CGGGAGGCCAGGAGGGCTCTGACTCTGTAAGTCGCTGAAGCTGACATGCCCTCA	705
PF		Qy	141	SerTDPArgIleArgIstyArgLeuSerGluAspAsnGlyTripleLeuY	159
XX	30-MAY-2000; 2000US-0208427P.	Db	706	TCCGGACGGATGAGACACTAGGATCCACTGCTGACAATGGCTGCTGTAC	762
PR					
XX	(LEACH M. D. (MEHR/.) MEHRABAN P. (TOPP/.) TOPPER J. N. (LAWD/.) LAW D.				
PA					
PI	Leach MD, Mehraban F, Conley PB, Topper JN, Law D;				
XX	DPI; 2002-625554/67.				
DR	P-PSDB; ABP64107.				
XX	PT				
PT	New polypeptide designated ORFX are present in human atherogenic cells and are useful to prevent and treat ORFX-associated disorders including cancer, allergy, wound healing or autoimmune, cardiovascular or inflammatory disease.				
XX	PS				
XX	Claim 2; SEQ ID NO 953; 78pp; English.				
CC	The present invention relates to novel human ORFX polypeptides and their coding sequences (ABP64641 and ABQ98194-ABQ99267). The sequences were discovered in human atherogenic cells, in particular in platelets and human umbilical vein endothelial cells (HUEBC) and are expressed in many other tissues as well. Atherogenic cells are cells which have the potential to develop atherosclerotic plaques. The ORFX polypeptides and nucleic acids are useful for treating or preventing a pathological condition associated with an ORFX-associated disorder, e.g. cancer, cardiovascular disease, allergy, autoimmune disease, wound healing, blood coagulation disorders or inflammatory disorders. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov/sequence.html?docID=20020082206				
XX	SQ				
XX	Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;				
XX	Alignment Scores:				
Qy	Pred. No.: 2.59e-63	Length:	763		
Score:	826.00	Matches:	158		
Percent Similarity:	99.37%	Conservative:	0		
Best Local Similarity:	99.37%	Mismatches:	1		
Query Match:	61.05%	Indels:	0		
DB:	6	Gaps:	0		
XX	US-09-939-853A-75 (1-261) x ABQ98670 (1-763)				
Qy	1 MetGlySerIeuProSerArgArgSerLeuProSerProSerLeuSerSerSerVal	20			
Db	286 ATGGAGATGTCGGAGAAATCTGCAAGCCAACTTGAGTCCTCCTC	345			
Qy	21 GlngLyngGlyProvalThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu	40			
Db	346 CAAGGCCAGGAACTGTGACCATGGAGCKAGAGCAAGGACACCGTGCCCTG	405			
Qy	41 GlySerPheProAlaGlyGlyProAlaGluLeuSerIeuArgLeuGlyGluProLeuThr	60			
Db	406 GGCAAGTTCCCCGGCACGTCAGATGGCTGAGCTGCTGAGCTGCTGAGCTGAC	465			
Qy	61 IleValSerGluAspGlyAspDPrTPrThrValLeuSerIeuValSerGlyArgGluTyr	80			
Db	466 ATGGCTCTGAGATGGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGAC	525			
Qy	81 AsnIleProSerValHisValCysIysValSerHisGlyTyrGluGlyLeuSer	100			
Db	526 AACATCCCAGGTCACAGTGGCTGATGGGGCTGATGGCTGAGC	585			
Qy	101 ArgGluIysAlaGluGluLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle	120			
XX	Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;				
XX	Alignment Scores:				

Pred. No.: 4.2e-55 Length: 603 PR 30-MAY-2000; 2000US-0208427P.  
 Score: 731.00 Matches: 151 XX  
 Percent Similarity: 82.15% Conservative: 4 PA (LEAC/)  
 Best Local Similarity: 80.32% MHRABAN M. D.  
 Query Match: 54.03% Mismatches: 21 PA (MEHR/)  
 DB: 5 Indels: 12 PA (CONL/)  
 Gaps: 4 DE (TOPP/)  
 US-09-939-853A-75 (1-261) × AAS74748 (1-603) Topper J.N. (LAW/)  
 Leach MD, Mehraban F., Conley PB, Topper J.N., Law D;  
 Qy 78 ArgGluGlyArgLeuSerValHisValGlyLysValSerHisGlyTrpLeuThrGlu 97 PI SEQ ID NO 1915; 78pp; English.  
 DR WPI; 2002-626554/67.  
 Db 61 CGCACGGGAACCTCCAGGAGTCGGTAC--- PA-P5DB; ABP64681.  
 PT New polypeptide designated ORFX are present in human atherosclerotic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 XX  
 Qy 98 GlyLeuSerArgGluLysAla----- GluGluLeuLeuLeuProGlyAsn 113 PT  
 Db 115 ---TTGCAACGGCTAAACGCTAACGCCCTAGAGCAGGAGAAGCTGAGCCTTGAG 171 PT  
 PT  
 Qy 114 ProGlyGlyAlaPhenLeuArgGluSerGlnThrArgArgGlySerTyrosLeuSer 133 PT  
 Db 172 ACCAACCTGTAGTCCTG---- CAAAGACATGGCTTACTCTGTCA 216 PT  
 PT  
 Qy 134 ValArgLeuSerArgProAlaSerTrpAsPArgLeuArgHistYArgLeuHisCysLeu 153 PT  
 Db 217 GTCCCGCTAGCGGCCTGATCCGGACGGATCACAGATCCACTGCTGCCT 276 PT  
 PT  
 Qy 154 AspAsnGlyTrpLeuTyrlSer-ProArgLeuThrPheProSerLeuGlnIleLeuVal 173 PT  
 Db 277 GACAATGGCTGGCTTACATCTCACCGCGCTCACCTTCCCCTACTCAGGCCCTGTG 336 PT  
 PT  
 Qy 174 AspHistYArgSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysValLeu 193 PT  
 Db 337 GACCATTAATCTGACCTGGGGATGACATCTGCTCCACTAAAGAGCCCTGTGCTCG 396 PT  
 PT  
 Qy 194 GlnArgAlaIgLyProLeuProLeuProLeuProValThrValGlnArgThr 213 PT  
 Db 397 CAGAGGGCTGGCCGCTCCCTGGAGGATAACCCPACTGTGACTGTGGAGGACA 456 PT  
 PT  
 Qy 214 ProLeuAsnTrpIleGluLeuAspSerSerLeuLeuPheSerGluLeuAlaIleArgGly 233 PT  
 Db 457 CCACTCAACTCGAAAGAGCTGAGCGTCCACAGGGAG 516 PT  
 PT  
 Qy 234 GluUserLeuSerGluGlyLeuArgGluUserSerPheTyrlSerLeuAsp 253 PT  
 Db 517 GAGTCCTTCTCAGTGAGGTCTGGGAGTCCCTAGCTGAATGAC 576 PT  
 PT  
 Qy 254 GluAlaValSerLeuAspAla 261 PT  
 Db 577 GAGGGCTGTCTTGGATGATGCC 600 PT  
 PT  
 RESULT 12 ABQ99151 ID ABQ99151 standard; DNA: 875 BP.  
 AC ABQ99151; XX  
 DT 04-NOV-2002 (First entry)  
 DE Human ORF58 coding sequence.  
 XX Homo sapiens.  
 KW Cyrostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antiinflammatory; gene therapy; human; ORFX; atherosogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 OS Homo sapiens.  
 XX  
 PN US2002082006-A1.  
 XX  
 PD 27-JUN-2002.  
 XX  
 PF 30-MAY-2001; 2001US-00867550.  
 XX  
 PR 250 SerIeuaAspGluIalaValSerLeuAspAspAla 251  
 DB 304 ACCCTGATGAGCAGGGTGTCTCTGGATGCC 339  
 ID ABX62975 standard; CDNA; 3756 BP.  
 ID ABX62975  
 RESULT 13 ABX62975





PR 02-OCT-2000; 2000US-0237295P  
 PR 02-OCT-2000; 2000US-0237316P  
 PR 03-OCT-2000; 2000US-023425P  
 PR 03-OCT-2000; 2000US-0237598P  
 PR 03-OCT-2000; 2000US-0237604P  
 PR 03-OCT-2000; 2000US-0237606P  
 PR 03-OCT-2000; 2000US-0237608P  
 PR 01-NOV-2000; 2000US-0244867P  
 PR 01-NOV-2000; 2000US-0245084P  
 XX  
 (AVAL-) AVALON PHARM.  
 PA  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 WPI; 2002-188264/24.  
 DR  
 XX  
 Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.  
 XX  
 PS Claim 1; SEQ ID NO 3526; 44P; English.  
 XX  
 The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour  
 XX  
 SQ Sequence 2665 BP; 736 A; 617 C; 689 G; 623 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 7.55e-33 Length: 2665  
 Score: 487.00 Matches: 101  
 Percent Similarity: 57.20% Conservative: 46  
 Best Local Similarity: 39.30% Mismatches: 94  
 Query Match: 35.99% Indexes: 16  
 DB: 6 Gaps: 3

US-09-939-853A-75 (1-261) × ABL65189 (1-2665)

Qy 5 ProSerArgGlySerLeuProSerProSerLeuSerSerValAlaThrIleLeuSerGlyGlyGlyGly 24  
 Db 24 CCAGGGAAAGAAAGAAATGGAAACAGGATGAATCCACCCCTGCCTGCCGAGAGG 83

Qy 25 ProValIthrMetGluAlaGluIargSerIyAlaValAlaLeuIysrI-PhePro 44  
 Db 84 CCCTTGCCAAACCGGAGGGACTGATGGACTCTCCGNGCTAAGTGTACTACCG 143

Qy 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyIuproLeuThrIleLeuSerGlu 64  
 Db 144 TCCTCCTCACATAGCCCCCATATTCCGGGAGGAACACTGCGTGATTCTCAT 203

Qy 65 AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyrosineProSer 84  
 Db 204 GAAGGGGGCTGGAAAGCTTATTCCTAGCTGAGGAGGATCACCCCTGGA 263

Qy 85 ValHisValGlyValSerHisGlyTipeLeuTyrglucIyLeuSerArgGluLysAla 104  
 Db 264 ATATGTTGGCCAGAGTTACCATGGCTGGCTGGTGGTTGAGGAGGAGAACGCC 323

Search completed: February 20, 2004, 09:56:47  
 Job time : 377 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:39:38 ; Search time 2518 Seconds  
 (without alignments)  
 3095.325 Million cell updates/sec

Title: US-09-939-853A-75  
 Perfect score: 1353  
 Sequence: 1 MGSLPSPRRKSLPSLSSV.....RESISFYISINDEAVSLLDA 261

Scoring table: BLOSUM62  
 Xgapop 10.0 , Xgapext 0.5  
 Ygapop 10.0 , Ygapext 0.5  
 Fgapop 6.0 , Fgapext 7.0  
 Delop 6.0 , Delext 7.0

Searched:

2713289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:

55026578

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

#### Command line parameters:

```
-Q=cgn2_1/USP0_spool.p/USP0_spool.p -APP QUERY.fasta_1.455
-DB=EST -QFMT=fastap -SFIX_P2N -MINMATCH=0.1 -LOCPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=15
-DOCALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGNN=15 -MODE=LOCAL
-OUTFMT=PL0 -NORM=Oct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US093953@CGN_1.1-423@runat_19022004_145339_24196 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=30 -WARN TIMEOUT=30 -THREADS=1 -XGAPPOP=10 -XGAPEXT=0.5 -FGAPPOP=6
-FGAEEXT=7 -YGAPOD=10 -IGAPEXT=0.5 -DELEXT=7
```

#### Database :

EST:\*

```
1: em_estba:*
2: em_estbn:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl1:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_eston:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fn:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_dro:*
25: em_gss_fod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
```

#### ALIGNMENTS

RESULT	1	BQ052308	1002 bp	mRNA	EST 29-MAR-2002
LOCUS		AGENCOURT	6868571 NIH MGC_106	Homo sapiens	linear cDNA clone IMAGE 5933542
DEFINITION					5' mRNA sequence.
ACCESSION		BQ052308			
VERSION		BQ052308.1			
KEYWORDS					EST.
SOURCE					
ORGANISM					Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE					1 (bases 1 to 1002)





Db	572 CGGGAGCCAGCAGGAGGCTTACTCTGTAGTCGGCTAGGCCCTGCA	631	AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Qy	141 SerIaspArgIleAspGlyTyrGlnLeuIleHisCysLeuAspGlyTyrIle	160	TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Db	632 TCCTGANGACCGATCAGAACCTACGGATGACTGGCTGTGACATC	691	JOURNAL	Nature 420, 563-573 (2002)
Qy	161 SerProArgLeuThrPheProSerLeuGlnAla-LeuValAspHistYerSerGluLeuAla	180	REFERENCE	6 (bases 1 to 2637)
Db	692 TCACCGCGCTCACCTGACTCCATGCCCTGAGCCCTGGTGCAGCTACTGGCTGGC	751	AUTHORS	Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, J., Hayashi, K., Hayashi, K., Hayashi, T., Hara, A., Hashizume, W., Hirono, T., Horio, T., Itoh, M., Itoh, M., Iwamoto, K., Hiraoaka, T., Kasukawa, T., Katsukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Ohno, M., Ohniso, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiragawa, A., Shirasaki, T., Sorabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku, A., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Miramatsu, M., and Hayashizaki, Y.
Qy	180 aAspAspIleCysCysLeu-LeuLysGluProCysValLeuGlnArgAlaGlyProLeuP	200	JOURNAL	Submitted (16 APR 2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-2 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome.gsc.riken.go.jp); URL: http://genome.gsc.riken.go.jp/, Tel: +81-45-503-9222, Fax: +81-45-503-9216
Db	752 GGATCACTATGCTGCTTCTACATCGGGACCCCTGCTGAAAGGTGaccGGCTCC	811	COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Sequencing Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Qy	200 roglIlysAspIle-ProLeuProValThrValGlnArgThrProLeuAspTrpLeuGlu	219	JOURNAL	Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Db	812 CTGGCAAAGTATAACCCATTACCCATTACCCATTACCTGCAAGAACCACTCATGCGGAG	871	COMMENT	Please visit our web site for further details.
Qy	220 LeuAspSerSerLeuLeu-LysSerGluLeuSerIleSerLeuLeuSer 238		JOURNAL	URL: http://fantom.gsc.riken.go.jp/
Db	872 CTGGGAGCTCCCTCTGGTTCTGAGCTGCCAGGGGGAGCTCTCTCATT	931	COMMENT	URL: http://fantom.gsc.riken.go.jp/
Qy	239 GluGly-LysArgGluSerLeuSerPhsThrIleSerLeu-AasnAspGluAla	255	JOURNAL	Location/Qualifiers
Db	932 GAGGGGTCTCCGGGAGTCCTGCTCTACATCAACCTGTATGACCAGGT	985	COMMENT	1. . 2637
RESULT	AK088672	2637 bp	FEATURES	/organism="Mus musculus"
LOCUS	AK088672	linear	source	/mol type="mRNA"
DEFINITION	Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone E430023D24 product: MODULATOR OF ANTIIGEN RECEPTOR SIGNALING MARS, full insert sequence.			/strain="NOD"
ACCESSION	AK088672			/db_xref="FANTOM DB:E430023D24"
VERSION	AK088672.1	GI:23353739		/ab_xref="MGI:2427569"
KEYWORDS	HTC; CAP trapper.			/clone_id="RIKEN full-length enriched mouse cDNA library"
SOURCE	Mus musculus (house mouse)			/dev_stage="2 days neonate"
ORGANISM	Bikaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			/dev_stage="358 . 1137"
REFERENCE	Carninci, P.			/note="unnamed protein product: MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (STR) AK088672, evidence: PASTY, purative!"
AUTHORS	Carninci, P. and Hayashizaki, Y.			/protein_id=BAC0495_1"
TITLE	High-efficiency full-length cDNA cloning			/codon_start_1
JOURNAL	Normal. Mol. Function. 303, 19-44 (1999)			/db_xref="GI:2635730"
MEDLINE	99-792953			/translation="MGSSSSRKEKTSSESPSSGGDPDPSMOPERHKVTAVALGSFPA GEQRSLSLRGEPHTISEGDWTVOSEVSGREHAFRGRGYSLSRLSRSPWDRTRHYRQLRGLWLYTS QKELLLPGNGGAFLIRESCOTRGRGYSLSRLSRSPWDRTRHYRQLRGLWLYTS PRTEPPSGLAHVPHYSELADGICCPREPCVQKLGPGLDTPPPVTPSSLNKK LDRSLLEPLAPGEASILSSEGURLESUSSYISLAEDPLDDA"
PUBMED	10349636			
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibusawa, K., Itoh, M., Konno, H., Ariyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishina, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwasa, M., Obara, E., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muranatsui, M., Inoue, Y., Kira, A., and Hayashizaki, Y.			
AUTHORS	RIKEN integrated sequence analysis system (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	1042159			
PUBMED	1042159			
REFERENCE	3 Sibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Ariyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishina, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwasa, M., Obara, E., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muranatsui, M., Inoue, Y., Kira, A., and Hayashizaki, Y.			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature 409, 685-690 (2001)			
REFERENCE	5			
Alignment Scores:				
Pred. No.:	4	8.96e-93	Length:	2637
Score:		1028.00	Matches:	208
Percent Similarity:		85.50%	Conservative:	16
Best Local Similarity:		79.39%	Mismatches:	34
Query Match:		75.98%	Index:	4
DB:		11	Gaps:	3

- |  |   |   |   |  |
|--|---|---|---|--|
| US-09-939-853A75 (1-261) x AK08672 (1-2637)  | 1 MetGlySerIleuProSerArgGlySerLeuProSerProSerLeuSerSerSerSerSerSer 20 | Qy  | 358 ATGGGAGTTTCACAGGGAAACCC--TCCGCCAACGCCAGCTCTCTGCT 414  | Db   |
| 21 GlyGlyGlyProValThrNetGluAlaGluSerLysAlaThrAlaValAlaLeu 40   | Qy  | 415 CGAACCGAACCCGNTTCATGCACCAGAACAGCACAGGTCAAGCTGGCCCTG 474   | Db  |  |
| 41 GlySerIheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuArgLeuGlyGluProLeuThr 60   | Qy  | 475 GGAGTTCCAGGTGAAGGCAAGACTCTGAACTCTGACTCAGTGAGCTGGGAGGCCGTCACC 534  | Db  |  |
| 61 IleValSerGluAspGlyAspPPTPPTrothrValLeuSerGluValSerGlyArgGluThr 80   | Qy  | 535 ATGATCTGAGATGGATGGATTTGGACAGTCAGTGAGCTAGGGAGAGACTAC 594   | Db  |  |
| 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrGlyLeuSer 100   | Qy  | 595 CACATGCCAACGAGTGTGTTGAGCTACGGCCCTGAG 654  | Db  |  |
| 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnPrGlyGlyAlaPheLeuIle 120   | Qy  | 655 CGCGGAAAGCCGAGGAACTACTCTGTTACCTGGAAACCCGGAGGGCCCTCCATC 714  | Db  |  |
| 121 ArgGluSerGlnThrArgGlySerTyrSerLeuSerArgProAla 140  | Qy  | 715 CGGGAGGCCAACGAGGAGGCTGCTGCTGACTACGGCCCTGACTACGGCCCTGCA 774  | Db  |  |
| 141 SerTPSPArgLysLeuArgHistYArgLeuHisCysLeuAspAsnGlyTriPleuYTrile 160  | Qy  | 775 TTTGGAACCGGATCAGACACTACAGATACTACGGCTCTTGACATGCTGGTGTACATC 834   | Db  |  |
| 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180   | Qy  | 835 TACCTGGCTTACCTTCCCCTACTCAGGCTGCTGAGATTACTCTGAGTAGCA 894   | Db  |  |
| 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGluArgAlaGlyProLeuPro 200   | Qy  | 895 GATGGCATCTGCTGTGCTCCCTCACTGGAGCCGCTGCTGCTGAGCTGGCCACTACCT 954   | Db  |  |
| 201 GlyLysAspDipeptideProLeuProValThrValGlnArgThrProLeuAsnTriPheGluLeu 220   | Qy  | 955 GCCTAAAGATAACCTCCACCTGACTGAGTGGCAACATCTACTAAATTGGAAAARGCTG 1014   | Db  |  |
| 221 AspSerSerLeuLeuPhsSerGluAla---AlaThrGlyGluSerLeuSerLeuSerGlu 239   | Qy  | 1015 GACCGCGGCCCTCTGTTGAAAGGACCTGCGAGTGGGAGCACTCTGTCAGTGAG 1074   | Db  |  |
| 240 GlyLeuLargGluSerLeuSerPheTyrIleSerLeuAsnPGluAlaValSerLeuAsp 259  | Qy  | 1075 GAGCTCCGAGTGGCTCCTCAGTCATCACAGCTGCTGAGCTGAGGAC----CCCTGGAT 1128  | Db  |  |
| 260 AspAla 261   | Qy  | 1129 GATGCT 1134  | Db  |  |
| RESULT 5<br>AK030877   | 2974 bp mRNA linear RTIC 18-SEP-2003                                  | LOCUS   | Mus musculus adult male thymus cDNA library, clone 5830437K0 product: MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence. | DEFINITION   |
| ACCESSION AK030877   | HTC CAP trapper.  | VERSION   | AK030877.1 GI:26326848  | KEYWORDS   |
| SOURCE Mus musculus (house mouse)  | ORGANISM Mus musculus   | REFERENCE Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Buteleostomi; Muridae; Murinae; Mus. Buxarota; Metazoa; Chordata; Scturognathici; Muridae; Murinae; Muridae; Murinae; Mus. Carnivora; P. Carninci, P. Hayashizaki, Y. High-efficiency full-length cDNA cloning J. Enzymol. 303, 19-44 (1999)  | FEATURES Location/Qualifiers  | COMMENT  |
| Carninci, P., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Murashita, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genomic Res. 10 (10), 1617-1630 (2000)   | REFERENCE   | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akizawa, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, M., Hazama, M., Nishimura, T., Harada, A., Yamamoto, R., Matsutomo, H., Sakaguchi, S., Ikezami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Isikawa, T., Ozawa, K., Tanaka, T., Matsuuwa, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kiris, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipicillary sequencer Genome Res. 10 (11), 1757-1771 (2000)  | REFERENCE   | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) |
| 3 Carninci, P., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Murashita, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genomic Res. 10 (10), 1617-1630 (2000) | REFERENCE   | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akizawa, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, M., Hazama, M., Nishimura, T., Harada, A., Yamamoto, R., Matsutomo, H., Sakaguchi, S., Ikezami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Isikawa, T., Ozawa, K., Tanaka, T., Matsuuwa, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kiris, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipicillary sequencer Genome Res. 10 (11), 1757-1771 (2000)  | REFERENCE   | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) |
| 4 Carninci, P., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Murashita, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genomic Res. 10 (10), 1617-1630 (2000) | REFERENCE   | Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Haga, A., Hashizume, W., Hara, A., Hirashima, T., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kojima, Y., Kawai, J., Kojima, T., Kondo, S., Kondo, H., Kouada, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Onno, M., Ohnato, N., Okazaki, Y., Saito, R., Saito, H., Sekai, C., Sekai, K., Sakakuma, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobabe, Y., Tagami, M., Togawa, A., Takahashi, F., Takanuki-Akahira, S., Takeeda, Y., Togami, M., Tomaru, T., Toyoda, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehito-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome.res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: +81-45-503-9222, Fax: +81-45-503-9216) CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://fantom.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers | REFERENCE   | 1. 2974  |
| 5 Carninci, P., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Murashita, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genomic Res. 10 (10), 1617-1630 (2000) | REFERENCE   | /mol_type="mRNA"<br>/strain="C57BL/6J"<br>/db_xref="FANTOM DB:5830437K10"<br>/db_xref="MGI:239406"  | REFERENCE   |  |





Gustincovich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawabata, Y., Kedzierski, R.M., King, P.A., Konigsmark, A., Kurockin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglioli, D.R., Maltais, L., Marchionni, L., McKenzie, J., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavon, W.J., Perera, G., Peelo, G., Petrosky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.I., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Santoni, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyrschka-Boris, A., Yano, Y., Yano, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sacchi, K., Shiraki, T., Waki, K., Iizawa, K., Arakawa, T., Fukuda, S., Harai, A., Hashizume, W., Imoto, K., Ishii, Y., Itoh, M., Kagawa, T., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Watersson, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analyses of the mouse transcriptome based on functional annotation of 60,770 full-length DNAs

Nature 420, 563-573 (2002)

22354483 MEDLINE  
12466851 PUBMED  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9116  
Email: genome-res@gsc.riken.go.jp/  
URL: <http://Genome.gsc.riken.go.jp/>  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashirume,W., Hayashida,K., Hirozane,T., Hori,F., Imoto,N., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohseto,N., Saito,R., Sakazume,N., Sano,H., Saeki,D., Sato,K., Shibata,K., Shiraki,T., Taganami,M., Takeda,Y., Waki,K., Watanabe,A., Muranatsu,M. and Hayashizaki,Y.

Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Retina RNA was provided by Dr. Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES	source	Location/Qualifiers
1..660		/organism="Mus musculus" /mol_type="mRNA" /db_xref=taxon:10090 /COne="A930009E21" /tissue_type="retina" /dev_stage="adult" /lab_host="DHIOB" /clone_id="RIKEN full-length enriched, adult retina" /note="Site 1: Sall; Site 2: BamH; cDNA library was prepared and sequenced in Mouse Genome Encyclopedi

Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGAGAGAGATCAAGAACCTTCTTTTTTTTNTVN 3'], cDNA was prepared by using trihaloless thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTGAGTAAATTATATCCCCCCCCCC 3']. cDNA was cleaved with XbaI and BamHI. Vector; a modified pBluescript KS(+) after bulk excision from Lambda PLC I. -Retina RNA was provided by Stefano Gustinovich, Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA02115, USA, whose assistance we gratefully acknowledge.



U.S. - 09 - 939 - 653A - 75 (1-261) x BG284179 (1-566)  
SUBJECTS: "Microfilm technologies". NOTE: NIH MRC LIBRARY.

ORIGIN

DB	Accession	Organism	Source	Location/Qualifiers	Sequence
Dy	US-09-939-853A-75				ThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerThrAspArg 144       ACCGTCGGCTCTACTCTGTCACTGCCCTCAGCCGGGG 63
db	4				MetGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerThrAspArg 144       ACCGTCGGCTCTACTCTGTCACTGCCCTCAGCCGGGG 63
Dy	145	IleArgHistYArgIleHisCysLeuAspPanglTIPLeuTyrlleSerProArgLeu 164 ATCGACAGCTACAGGATCCACTGCCCTGACAATGGCTGGCTACATCTACCGGCCCTC 123			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 44 Score: 3.35E-58 Pred. No.: 680.50 Percent Similarity: 73.13% Best Local Similarity: 71.81% Query Match: 50.30% DB: 12
Dy	165	ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCys 184 ACCTTCCTCTACTCCGGCCCTGGTGAACATTACTCTGAGCTGGGATGACATCTGC 183			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
Dy	185	CysLeuLeuLysCysLeuProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204 TGCTACTAAAGGCCCTGTTCCAGGGCTGCCGCTCCCTGGRAGGATA 243			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
Dy	184	ProLeuProValThrValGlnArgThrProLeuAsnTrpIleGluLeuAspSerSerLeu 224 CTGTTCTGAACCTGCCCTCAGGAGCTCTCTGAGCTGGAGAGTGGAGAGTGGAGCTTC 303			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
Dy	205	LeuPheSerGluAlaAlaThrGlyGluGluSerIleLeuSerGluGlyLeuArgGluSer 244 CCTCTACTGACTGAGCAGAACCTCACTGGAGGAGCACCTAACCTGAAGAGTGGAGAGTGGAGCTTC 363			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
Dy	225	IeuSerPheTyrSerLeuAsnArgGluSerIleLeuSerGluGlyLeuArgGluSer 244 Q054265 261			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
Dy	304	IeuSerPheTyrSerLeuAsnArgGluSerIleLeuSerGluGlyLeuArgGluSer 244 CTCAGCTCTACATAGCTGATGAGGGATGTCAGGAGCTGCTCTGATGCC 414			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
Dy	245	LeuSerPheTyrSerLeuAsnArgGluSerIleLeuSerGluGlyLeuArgGluSer 244 B0054265 261			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
Dy	364	IeuSerPheTyrSerLeuAsnArgGluSerIleLeuSerGluGlyLeuArgGluSer 244 B0054265 261			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
Dy	365	RESULT 10			Length: 986 Matches: 163 Conservative: 3 Mismatch: 17 Gaps: 3
LOCUS	B0054265	AGENCOURT 6830248 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936339	mRNA	Linear EST 29-MAR-2002	Length: 986 bp mRNA linear EST 29-MAR-2002
DEFINITION		5', mRNA Sequence.			Length: 986 bp mRNA linear EST 29-MAR-2002
ACCESSION	B0054265				Length: 986 bp mRNA linear EST 29-MAR-2002
KEYWORDS	B0054265.1	GI:19813605			Length: 986 bp mRNA linear EST 29-MAR-2002
SOURCE	Homo sapiens (human)				Length: 986 bp mRNA linear EST 29-MAR-2002
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. National Institutes of Health, Mammalian Gene Collection (MGC)				Length: 986 bp mRNA linear EST 29-MAR-2002
REFERENCE	NIH-MGC URL: http://mgc.ncbi.nih.gov/				Length: 986 bp mRNA linear EST 29-MAR-2002
AUTHORS	Unpublished (1999)				Length: 986 bp mRNA linear EST 29-MAR-2002
TITLE	Contact: Robert Straussberg, Ph.D.				Length: 986 bp mRNA linear EST 29-MAR-2002
JOURNAL	Email: csapbsr@mail.nih.gov				Length: 986 bp mRNA linear EST 29-MAR-2002
COMMENT	Tissue Procurement: Dr. Daniel McVicar, DBS/NCI				Length: 986 bp mRNA linear EST 29-MAR-2002
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium				Length: 986 bp mRNA linear EST 29-MAR-2002
	DNA Sequencing by: Agencourt Bioscience Corporation				Length: 986 bp mRNA linear EST 29-MAR-2002
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNRL at:				Length: 986 bp mRNA linear EST 29-MAR-2002
	http://image.llnl.gov/Plate: LLNL2125 row: i column: 12				Length: 986 bp mRNA linear EST 29-MAR-2002
	High quality sequence stop: 515.				Length: 986 bp mRNA linear EST 29-MAR-2002
FEATURES	source				Length: 986 bp mRNA linear EST 29-MAR-2002
	/mol type="mRNA"				Length: 986 bp mRNA linear EST 29-MAR-2002
	/db_xref="taxon:9606"				Length: 986 bp mRNA linear EST 29-MAR-2002
	/clones="IMAGE:5936339"				Length: 986 bp mRNA linear EST 29-MAR-2002
	/tissue_type="natural killer cells, cell line"				Length: 986 bp mRNA linear EST 29-MAR-2002
	/lab_host="DH10B (phage-resistant)"				Length: 986 bp mRNA linear EST 29-MAR-2002
	/clone_lib="NIH MGC_106"				Length: 986 bp mRNA linear EST 29-MAR-2002
	/note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into BCR1/XhoI sites using the following 5' adapter: GCACAGG (G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZapBluescript synthesis kit (Stratagene) and				Length: 986 bp mRNA linear EST 29-MAR-2002
RESULT	11				Length: 986 bp mRNA linear EST 29-MAR-2002
	B0944126				Length: 986 bp mRNA linear EST 29-MAR-2002
LOCUS					Length: 986 bp mRNA linear EST 29-MAR-2002
DEFINITION					Length: 986 bp mRNA linear EST 29-MAR-2002
IMAGE:6728350 5'					Length: 986 bp mRNA linear EST 29-MAR-2002
B0944126					Length: 986 bp mRNA linear EST 29-MAR-2002
ACCESSION					Length: 986 bp mRNA linear EST 29-MAR-2002
VERSION					Length: 986 bp mRNA linear EST 29-MAR-2002
KEYWORDS					Length: 986 bp mRNA linear EST 29-MAR-2002
SOURCE	Homo sapiens (human)				Length: 986 bp mRNA linear EST 29-MAR-2002
ORGANISM	Homo sapiens				Length: 986 bp mRNA linear EST 29-MAR-2002





Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sugioke, Y., Suzuki, D.H., Tagami, M., Takahashi, F.,  
 Takeuda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTS (Arakawa, T., et al. 2001)

Contact: Yoshinori Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute of Physical and Chemical Research (RIKEN)  
 Tel: 81-45-503-9216  
 Fax: 81-45-503-9216  
 Email: genome-res@riken.go.jp/  
<http://genome.gsc.riken.go.jp/>  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. *Genome Res.* 10 (10), 1617-1630 (2000).  
 wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Ohara, E.,  
 Matsukubo, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsuuwa, S., Kawai, J., Okazaki, Y., Muranatsu, M., Inoue, Y., Kira, A.,  
 and Hayashizaki, Y.  
 RIKEN Integrated sequence analysis (RISA) system -384 format  
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
 10 (11), 1757-1771 (2000).  
 Konno H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 mouse full-length cDNA library. *Genome Res.* 11 (2), 281-289 (2001).  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamamoto, I.,  
 Alizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J.,  
 Shibata, K., and Hayashizaki, Y.  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001).  
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
 further details.

e mouse tissues.

Location/Qualifiers

- 1. .660
- /organism="Mus musculus"
- /mol\_type="mRNA"
- /db\_xref="taxon:10090"
- /clone="A33007R07"
- /tissue\_type="thymus"
- /dev\_stage="0 day neonate"
- /lab\_host="DH10B"
- /clone\_lib="RIKEN full-length enriched, 0 day neonate thymus"

/note="Site 1: SaliI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5, GAGAGAGAGGATCCAAAGGCTCTTCTTTTTTTVN 3']". cDNA was prepared by using trizolose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGGATTCCTGAGTTAATTAAATTAAATCCCCCCCCCCCC 3']. cDNA was cleaved with XbaI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FRT 11

ores:	6.67e-46	Length:	660
clarity:	55.0	Matches:	114
	86.52%	Conservative:	8

Location/Qualifiers  
1. .597  
organism="Homo sapiens"  
mol\_type="mRNA"  
db\_xref="taxon:9606"  
name="2D"

/clone\_lib="pool\_YT\_lib\_v\_SPD"

ORIGIN

	Alignment Scores:	Length:
Pred. No.:	7.98e-46	597
Score:	555.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	41.02%	Indels:
DB:	9	Gaps:

US-09-939-853A-75 (1-261) x AL844311 (1-597)

```

Qy 155 AsnGlyTyrLeuTyrIleSerProArgLeuThrPheProSerLeuGinAlaLeuValAsp 174
Db 13 ArgGCTGCGCTGACATCTCACAGCTAACCTCCCTACTCCAGGCCCTGGGAC 72
Qy 175 HistYrSerGluLeuAlaAspAspIleCysLeuLeuLysGluProCysValLeuGln 194
Db 73 CATTACTCAGGTCAGCATCAGCTGCAGTCAGCTACTCAAGAGCCCTGTCCTGAG 132
Qy 195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPro 214
Db 133 AGGGTGGCCGCTCCCTGCAAGGATAAACCTTACCTGTTGACTTGCGCAAGGAAACCA 192
Qy 215 LeuAsnTyrIleGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGlu 234
Db 193 CTCAACTGAAAGAGCTGGACAGTCCCTCAGGTTGAAAGCTGCAACGGGGAG 252
Qy 235 SerLeuLeuSerGluLeuArgGluLeuSerPheTyrIleSerLeuAsnAspGlu 254
Db 253 TCTCTTCTCAAGTGGGCTCCGGAGTCAGCTCTACATGCCATGAGCAG 312
Qy 255 AlaValSerLeuAspAspAla 261
Db 313 GCTGCTCTGTTGATGAGSC 333

```

Search completed: February 20, 2004, 11:30:54  
 Job time : 251 secs

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_P2n model

Run on: February 20, 2004, 09:40:13 ; Search time 89 Seconds  
(without alignments)  
1627.440 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLPSRRKSLPSPSSV.....RESLSFVISLNDAVSLLDA 261

Scoring table: BLOSUM62

Xgapext	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delext	6.0	Deletxt	7.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line Parameters:  
-MODEL=frame+P2n.model -DEV=xlp  
-Q=/USPTO\_spool\_P/US93853/runat\_19022004\_14540\_24508/app\_query.fasta\_1.455  
-DB=Issued\_Patents\_NA -QEMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=6its -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human0.cdt  
-LIST=45 -DOCALIGN=200 -THR SCORE=90CT -THR MAY=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTEMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFILE=BLOCK100 -LOGFILE  
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOB=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/pctodata/2/ina/5A\_COMB\_seq:  
2: /cgn2\_6/pctodata/2/ina/5B\_COMB\_seq:  
3: /cgn2\_6/pctodata/2/ina/6A\_COMB\_seq:  
4: /cgn2\_6/pctodata/2/ina/6B\_COMB\_seq:  
5: /cgn2\_6/pctodata/2/ina/PC10S.COMB.seq:  
6: /cgn2\_6/pctodata/2/ina/backfile1.seq:  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution. and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	370.5	27.4	2298	4	US-09-023-655-1158 Sequence 1158, Ap
2	360.5	26.6	2015	4	US-09-023-655-1105 Sequence 1105, Ap
3	340	25.1	2129	4	US-09-016-434-1452 Sequence 1452, Ap
4	323	23.9	2435	4	US-09-023-655-11313 Sequence 1313, Ap
5	320	23.7	2647	4	US-09-220-132-77 Sequence 77, Appl
6	320	23.7	2647	5	PCT-US93-06251-77 Sequence 77, Appl
7	315.5	23.3	4517	4	US-09-470-881-7 Sequence 7, Appl
8	313.5	23.3	4517	5	PCT-US93-06251-83 Sequence 83, Appl
9	313.5	23.2	1491	2	US-09-006-675-1 Sequence 1, Appl
10	313.5	23.2	1491	3	US-09-228-605-1 Sequence 1, Appl
11	311.5	23.1	2354	4	US-09-023-655-1080 Sequence 1080, Ap
12	289	21.4	1759	4	US-09-470-881-2 Sequence 2, Appl

## ALIGNMENTS

## RESULT

## 1

/ Sequence 1158, Application US/09023655

Patent No. 6607819

GENERAL INFORMATION:

APPLICANT: Jeffrey J. Seihamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

NUMBER OF SEQUENCES: EXPRESSION

CITY: PALO ALTO

STATE: CALIFORNIA.

COUNTRY: USA

ZIP: 94304

COMPUTER AVAILABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023, 655

FILING DATE: HEREWITH

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37, 071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1158:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2298 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: 9187268  
 US-09-023-655-1158

Alignment Scores:

Pred. No. :	9.15e-30	Length:	2298
Score:	370.50	Matches:	80
Percent Similarity:	57.71%	Conservative:	36
Best Local Similarity:	39.80%	Mismatches:	76
Query Match:	27.38%	Indels:	9
DB:	4	Gaps:	3

US-09-939-853a-75 (1-261) x US-09-023-655-1158 (1-2298)

Qy    6 SerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGlyPro 25  
 Score:    370.50    Length:    2298  
 Db    409 TCCATAAACGAAAGGCCAGTTCAGAA---TCTCAGCTTTACCTGGAGAGGTT 465  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    26 ValThrMetIluAlaGluArgSerTysAlaThrIluAlaLeuGlySerPheProAla 45  
 Score:    370.50    Length:    2298  
 Db    466 CAAATTAAGATCCAGGAAACAGGAACATTGTGGAGCCCTGTTACCCPATGATGGC 525  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    46 GlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGluAsp 65  
 Score:    370.50    Length:    2298  
 Db    526 ATCCACCGAACGACTTGTTTCAAGAAGGAGAGCATGAAAGTCCTGAGGAGCAT 585  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    66 GlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyrosAsnLeuProSerVal 85  
 Score:    370.50    Length:    2298  
 Db    586 GGAGATGGGAAAGCAAAGTCCTTTAACAAAAGAACGCTTATCCCCAGAAC 645  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    86 HisValGlyLysVal-----SerHisGlyTrpLeuTyrGluGlyLeuSerArg 101  
 Score:    370.50    Length:    2298  
 Db    646 TATGTCGCCAAACTAACCTTAGAAAGAGTGTTCAAAGATAAACAGG 705  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    102 GluValGluGluLeuLeuLeuProl1AsnProGlyGlyAlaPheLeuIleArg 121  
 Score:    370.50    Length:    2298  
 Db    706 AAGGGCAGAAAGCCAGCTTTGCACCGAAATAGCCGCTGAGCTTCTTATTAGA 765  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    122 GluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSer 141  
 Score:    370.50    Length:    2298  
 Db    766 GAAATGAAACATTAAAGGAGCTCTCTGCTGTCAGAGACTTGTGCAT 825  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    142 TrpAspArgIleArgHistYrrArgIleHisSerLeuAspArgIleHisSerLeu 161  
 Score:    370.50    Length:    2298  
 Db    826 GGTGATGTTATTAGACTACAATTAGAAGTGTGGATAATGGGGTATACATCT 885  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    162 ProArgLeuThrPheProSerLeuLeuLeuGlyAlaGlyProLeuProGly 201  
 Score:    370.50    Length:    2298  
 Db    886 COACGATCATCTTCCCTGTTACAGCAGATTAACAAAGCAGGAGAT 945  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    182 AspIleCysLeuLeuLeuGlySerLeuGlyGlnArgAlaGlyProLeuProGly 201  
 Score:    370.50    Length:    2298  
 Db    946 GSCTTGGCAGAAAGATTGGGAAAGGCTTGTATT-----AGTCCAGCCACAG 993  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Qy    202 Lys 202  
 Score:    370.50    Length:    2298  
 Db    994 ARG 996  
 Percent Similarity:    57.71%  
 Best Local Similarity:    39.80%  
 Query Match:    27.38%  
 DB:    4

Alignment Scores:

Pred. No. :	8.84e-29	Length:	2015
Score:	360.50	Matches:	77
Percent Similarity:	58.36%	Conservative:	31
Best Local Similarity:	41.62%	Mismatches:	70
Query Match:	26.64%	Indels:	7
DB:	4	Gaps:	2

US-09-939-853a-75 (1-261) x US-09-023-655-1105 (1-2015)

Qy    12 ProSerProSerLeuSerSerValGlnGlyGlnGlyProValThrMetGluAlaGlu 31  
 Score:    286 CGGGGCGCTTATAGCCACAAAGCAGCAGCTTACGAGGAGGAGGAGGAGC 336  
 Percent Similarity:    58.36%  
 Best Local Similarity:    41.62%  
 Query Match:    26.64%  
 DB:    4

Qy    32 ArgSerLysAlaThrAlaValAlaValAlaLysGlySerPheProAlaGluLeu 51  
 Score:    337 TCTGAGGACATCATCGTGGITGCCCTGATGATCACGAGCCATTCAAGAAGACCTC 396  
 Percent Similarity:    58.36%  
 Best Local Similarity:    41.62%  
 Query Match:    26.64%  
 DB:    4

Qy    52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspPipTrpThrVal 71  
 Score:    286 CGGGGCGCTTATAGCCACAAAGCAGCAGCTTACGAGGAGGAGGAGGAGC 336  
 Percent Similarity:    58.36%  
 Best Local Similarity:    41.62%  
 Query Match:    26.64%  
 DB:    4

Qy    72 LeuSerLeuValSerIleGlyArgIleProSerValHisValGlyIleVal 90  
 Score:    286 CGGGGCGCTTATAGCCACAAAGCAGCAGCTTACGAGGAGGAGGAGGAGC 336  
 Percent Similarity:    58.36%  
 Best Local Similarity:    41.62%  
 Query Match:    26.64%  
 DB:    4

RESULT 2  
 Sequence 1105, Application US/09023655  
 Patent No. 6607879  
 GENERAL INFORMATION:  
 APPLICANT: Cocks, Benjamin G.

Qy 108 LeuLeuLeuProGlyAsnProGlyGlyAlaAlaLeuArgLysSerGlnThrArgArg 127  
 Db 577 CTGGTGGCTGCCGAAACATGGCTGGCTTCATGATCCGGATAGCGAACCTCTAA 636  
 Qy 128 GlySerTyrosLeuSerValArgLeuSerArgProLeuSerTrpAspSerArgHis 147  
 Db 637 GGAGCTACTCTTGTCTGGAAACTGAGCTACGGCCCTGGCAGGGATAACGTTGAACT 696  
 Qy 148 TyroGlyIleHisCysLeuAspAsnGlyTrpLeuTyrosLeuSerProArgLeuThrPhePro 167  
 Db 697 TACAGATCGGACCCCTGGACAGGGGGCTTACATACTCCCCGAAGGACCTTCAGC 756  
 Qy 168 SerIleGlnAlanLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysIleLeu 187  
 Db 757 ACTTGACAGGTGGTGGACCATACAGGGAAACGAGGCTCTGCCAGAAACTG 816  
 Qy 188 LysGluProCysVal 192  
 Db 817 TCGGTGCCCTGCACTG 831

RESULT 3

US-09-016-434-1452  
 Sequence 1452, Application US/09016434  
 ; Patent No. 650038

; GENERAL INFORMATION:  
 ; APPLICANT: Janice Au-Young  
 ; APPLICANT: Jeffrey J. Seihamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 ; NUMBER OF SEQUENCES: 14 90

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/016,434  
 ; FILING DATE: HEREWITH  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeiller, Karen J.  
 ; REGISTRATION NUMBER: 37,071  
 ; REFERENCE/DOCKET NUMBER: PA-0002 US

; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 855-0555  
 ; TELEFAX: (650) 845-4166  
 ; INFORMATION FOR SEQ ID NO: 1452:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 229 base Pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GENBANK  
 ; CLONE: 9775207

US-09-016-434-1452

Alignment Scores:	
Pred. No. :	1.52e-26
Score:	340.00
Percent Similarity:	55.0%
Best Local Similarity:	40.56%
Length:	2129
Matches:	73
Conservative:	71
Mismatches:	71

```

Query Match:      25.13%    Indels:   10
DB:                4          Gaps:     2
US-09-939-853A-75 (1-261) x US-09-016-434-1452 (1-2129)

Qy   26 ValThrMetGluAlaGluArgSerLysAlaThr-----.
Db   204 GRACCTCGAACTCCATGCCGCTTCCACTGCAAG
Qy   40 LeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuP-----.
Db   264 CTCACACCTTAATAGCCCTCTCAGCAGGAGCTGGCTTC
Qy   60 ThrIleValSerGluAspGlyAspPTPTrpThrValLeuSer-----.
Db   324 CGGATCCCTGGAGGAGGGCGGAGGTGGAGGGCAGTGCTCC
Qy   80 TyrAsnIleProSerValIleValGlyLysValSer-----.
Db   384 GGTTTCATCCCCCTCAATTGGCC2AAGCCAAAGCCCTGG
Qy   96 TyrGluGlyLeuSerArgGluLysAlaGluGluLeuLeuLys-----.
Db   444 TCTAAGAACCTGTGCCGAAGGGCGGAGGCCAGCTCTGG
Qy   116 GlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySer-----.
Db   504 GGCCTCCCTCTCATCCGGAGAACGGAGAACGCCAGGGATCGG
Qy   136 LeuSerArgProMlaSerTriPaspArgIleArgHistYArgIle-----.
Db   564 GACTTGCACCAAGACCAAGGAGGGAGGTGGTGAACATPACAAGG
Qy   156 GlyTrpLeuTyrlIleSerProArgLeuThrPheProSerLeu-----.
Db   624 GTGGGCTCTCATCTCCCTCGATACTCTTCCCCTGGATCCTGG
Qy   176 TyrSerGluLeuAlaAspAspIleCysCysLysLeuLysGlu-----.
Db   684 TAGCCATGCTCAGATGGGTGTCACGGTTGAGCCGCC
RESULT 4
US-09-023-655-1313
; Sequence 1313, Application US/09023655
; Patent No. 6609879
; GENERAL INFORMATION:
;   APPLICANT: Cocks, Benjamin G.
;   APPLICANT: Susan G. Stuart
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF
;   TITLE OF INVENTION: EXPRESSION
;   NUMBER OF SEQUENCES: 1508
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;     STREET: 3174 PORTER DRIVE
;     CITY: PALO ALTO
;     STATE: CALIFORNIA
;     COUNTRY: USA
;     ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOSS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09-023, 655
;   FILING DATE: HERENTH
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;     CLASSIFICATION:
;     AGENT INFORMATION:

```

NAME: Zeller, Karen J.  
 REFERENCE/DOCKET NUMBER: 37,071  
 TELEPHONE: (650) 845-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 131:3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2435 base pairs  
 TYPE: nucleic acid  
 STANDINGNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: 9338227  
 US-09-023-655-1313

Alignment Scores:  
 Pred. No.: 1.24e-24 Length: 2435  
 Score: 323.00 Matches: 81  
 Percent Similarity: 51.74% Conservative: 23  
 Best Local Similarity: 40.30% Mismatches: 76  
 Query Match: 23.87% Indels: 22  
 DB: 4 Gaps: 3

US-09-939-853a-75 (1-261) x US-09-023-655-1313 (1-2435)  
 Qy 12 ProSerProSerLeuSerSerSerValGlnGly-----22  
 Db 507 CAATCCCAACTACACAACCTCCAGCAGGCCGGGGCAAGGACTCACCGCTTTGAG 566

Qy 23 -----GlnGlyProValThrMetGluAlaGluArgSerLysAla 35  
 Db 567 GTGTGACTCTCGTCTCATCGGAGACTTGCTACGAGAGGAACTGGAGGAGTGGAGTGG-ACA 625

Qy 36 ThrAlaValAlaLeuGlySerSerProAlaGlyGlyProAlaGluLeuSerLeuArgLeu 55  
 Db 626 CTCTTGTGGCCCTTATGACTATGAGCTGAGCTGGTTTCACAA 685

Qy 56 GlyGlyProLeuThrIleVal---SerGluAspGlyAspTyrTrpThrValLeuSerGlu 74  
 Db 686 GGAGAAAAAATTCAAATATGAAACAGCTCGGAAGGGAGATTGGTGGAGGCCGCTCCCTG 745

Qy 75 ValSerGlyArgGlyTyramileProSerValIleHisvalGlyLysVal---90  
 Db 746 ACAACTGGAGAGACGGTTACATCCAGGAATPATGGCTCAGTGCACCTATCCAG 805

Qy 91 SerHisGlyTrpLeuThrGluGlyLeuSerArgGlyIleLeuLeuLeuLeu 110  
 Db 806 GCAGAGAGGTGACTTGAAAATTGGCGGAAAGATGCTGAGCAGCTATTGTC 865

Qy 111 ProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerLysArgGlySertYR 130  
 Db 866 TTGGAAACCAAGAGTACCTTCTATGGAGAGTAAACCAAGGTTCPAT 925

Qy 131 SerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIle 150  
 Db 926 TCACCTTCTATCGTGTATTGGGATGATGAAAGACATGTCAAATAAATT 985

Qy 151 HisCysLeuAspAsnGlyTrpLeuThrPheProSerLeuGln 170  
 Db 986 CGCAAActTGCAATGTTGCGATACTACATGCACTGGCCAGTTGAAACACTTCAG 1045

Qy 171 AlaLeuValAspHistYrsSerGluLeuAlaAspAspIleCysLeuLeuLysGluPro 190  
 Db 1046 CAGCTTGATACAACTTAATCTAGAGAGCTGAGGTCAGGTTCTCGCCCTAGAGTGTCCC 1105

Qy 191 Cys 191  
 Db 1106 TGT 1108

RESULT 5  
 US-09-220-132-77

---

i Sequence 77, Application US/09220132  
 i Patent No. 6506607  
 i GENERAL INFORMATION:  
 i APPLICANT: Shylian, Andrew W.  
 i METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
 i TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT OF PROSTATE CANCER  
 i TREATMENT  
 i FILE REFERENCE: 07334-074001  
 i CURRENT APPLICATION NUMBER: US/09/220,132  
 i CURRENT FILING DATE: 1998-12-23  
 i PRIORITY APPLICATION NUMBER: US 60/079,303  
 i PRIORITY FILING DATE: 1998-03-25  
 i PRIORITY APPLICATION NUMBER: US 60/068,821  
 i PRIORITY FILING DATE: 1997-12-24  
 i NUMBER OF SEQ ID NOS: 191  
 i SOFTWARE: FastSSQ for Windows Version 4.0  
 i SEQ ID NO: 77  
 i LENGTH: 2647  
 i TYPE: DNA  
 i ORGANISM: Homo sapiens  
 i US-09-220-132-77

Alignment Scores:  
 Pred. No.: 2.97e-24 Length: 2647  
 Score: 320.00 Matches: 80  
 Percent Similarity: 51.74% Conservative: 24  
 Best Local Similarity: 39.80% Mismatches: 76  
 Query Match: 23.65% Indels: 22  
 DB: 4 Gaps: 3

US-09-939-853a-75 (1-261) x US-09-220-132-77 (1-2647)

Qy 12 ProSerProSerLeuSerSerSerValGlnGly-----22  
 Db 716 CCATCCCCAAACTACACAAACTCCACGACCGGGGGCAAGGACTCACCGCTTTGAG 775

Qy 23 -----GlnGlyProValThrMetGluAlaGluArgSerLysAla 35  
 Db 776 GTGTGACTCTCTCGTCTCATCGGGGAACTTGCTACGAGAGGAACTGGAGTGG-ACA 834

Qy 36 ThrAlaValAlaLeuGlySerSerProAlaGlyGlyProAlaGluLeuSerLeuArgLeu 55  
 Db 835 CTCTTGTGGCCCTTATGACTATGAGGAGCCGAGGAGCTGACTTTTCACAAA 894

Qy 56 GlyGlyProLeuThrIleVal---SerGluAspGlyAspTyrTrpThrValLeuSerGlu 74  
 Db 895 GGAGAGAAATTCAAATATGAAACAGCTCGGAAGGGAGATTGGTGGAGGCCGCTCCCTG 954

Qy 75 ValSerGlyArgGlyTyramileProSerValIleHisvalGlyLysVal---90  
 Db 955 ACAACTGGAGAGACGGTTACATCCAGGAATPATGGCTCAGTGCACCTATCCAG 1014

Qy 91 SerHisGlyTrpLeuThrGluGlyLeuSerArgGlyIleLeuLeuLeu 110  
 Db 1015 GAGAGAGGTTGACTTGAAAATTGGGATGATAAGGAGCATGTCAGCTATTGTC 1074

Qy 111 ProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerLysArgGlySertYR 130  
 Db 1075 TTGGAAACCAAGAGTACCTTCTATGGGATGATGAAAGACATGTCAAATAAATT 1134

Qy 131 SerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIle 150  
 Db 1135 TCACCTTCTATCGTGTATTGGGATGATAAGGAGCATGTCAGCTATTGAAACACTTCAG 1194

Qy 151 HisCysBleuAspAsnGlyTrpLeuThrPheProSerLeuGln 170  
 Db 1195 CGCAAActTGCAATGTTGCGATACTACATGCACTGGCCAGTTGAAACACTTCAG 1254

Qy 171 AlaLeuValAspHistYrsSerGluLeuAlaAspAspIleCysLeuLeuLysGluPro 190  
 Db 1255 CAGCTTGATACAACTTAATCTAGAGAGCTGAGGTCAGGTTCTCGCCCTAGAGTGTCCC 1314

Qy 191 Cys 191



SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US93/06251  
 FILING DATE: 19930630  
 CLASSIFICATION:  
 NAME: DiGiilio, Frank S.  
 REGISTRATION NUMBER: 31,346  
 REFERENCE/DOCKET NUMBER: 8586  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 516-742-4343  
 TELEFAX: 516-742-4366  
 TELEREX: 230 901 SANS UR  
 INFORMATION FOR SEQ ID NO: 83:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4517 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 PCT-US93-00251-83

Alignment Scores:  
 Pred. No.: 2.05e-23 Length: 4517  
 Score: 315.50 Matches: 92  
 Percent Similarity: 43.22% Conservative: 45  
 Best Local Similarity: 29.02% Mismatches: 113  
 Query Match: 23.32% Indels: 67  
 DB: 5 Gaps: 9

US-09-939-853A-75 (1-261) x PCT-US93-06251-83 (1-4517)

Qy 2 GlySerLeuProSerArgArgLysSerLeuProSerLeuSerSerVal--- 20  
 Db 304 GGAGCAGAACCCACTACAGTCATCATGTCGTCATCTCAGCAAAGGGAAACAGAGTT 363

Qy 21 GluArgSerLysAlaThrAlaValAlaLeuLysSerProAlaGluLeuSerLeu 30  
 Db 364 AATTCAAGCAGCTTCATGACACCATTCAGGGATCTCAGGTTAACGCCTTTGGA 423

Qy 31 GluArgSerLysAlaThrAlaValAlaLeuLysSerProAlaGluLeuSerLeu 46  
 Db 424 GGTCATCTCTCATTTCACTTCAGTGGTCCAAAGTTCACTATCTGTGGTTAACAGTTGGT 483

Qy 47 GluProAlaGluLeuSerLeu 53

Db 484 GTTACIATATTGTGGCTTATATGATTAAGTGAACTACAGAGCTTCTCTTT 543

Qy 54 ArgLeuGlyGluProLeuThrIleValSerGlu---AspGlyAspTrpThrValLeu 72

Db 544 AAGAAGGCTGAAATTCAGGATGTTCAAGACTAGGTTAAAGATGTTAACAGGAAAGAGATA 603

Qy 73 SerGluValSerGluGlyGluLeuSerLeu 90

Db 604 TCAATCGCTACAGGAAGTAATGTCGCCGAAATTATGCCGAGATTCAGGCCAGATTC 663

Qy 91 SerHisGlyTrpLeutyrGluGlyLeuSerArgGluLysAlaGluGluLeu 108  
 Db 664 ATTTCAGGCGAGAAAGTAATGTCGCCGAAATGTCGCCGAAAGATCTGAAAGATTCTT 723

GENERAL INFORMATION:  
 APPLICANT: Wickstrom, Eric and Rife, Jason P.  
 TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
 NUMBER OF SEQUENCES: 93  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SCULLY, MURPHY & PRESSER  
 STREET: 400 Garden City Plaza  
 CITY: Garden City  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 11530  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS

RESULT 8  
 PCT-US93-06251-83  
 Sequence 83, Application PC/TUS9306251  
 GENERAL INFORMATION:  
 APPLICANT: Wickstrom, Eric and Rife, Jason P.  
 TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
 NUMBER OF SEQUENCES: 93  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SCULLY, MURPHY & PRESSER  
 STREET: 400 Garden City Plaza  
 CITY: Garden City  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 11530  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS

Qy 147 HistYrArgIleHisCysLeuAspArgGlyTrpLeutyrIleSerProArg 163  
 Db 829 AATGTGAAACATACAATATTAGGAACACTGCAATCTAGTAAAGACTAACTATATAACACAGA 888

Qy 164 LeuThrPheProSerLeuGlnAlaLeuValAspHistYrSerGluLeuValAsp 183

Db 889 GCACAAATTGATACTCTGAAATTGAAACACTACAGACATCTGATGGTTA 948

Qy 184 CysCysLeuLeuLysGluProCys-----ValLeuGln 194

Db 949 TGCCACAGTTGACAACACTGGTCAACCTGAAACCTAGACTAAAGCTTAGCAAAA 1008

Qy 195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArg--- 212

Db 1009 GATGCTGGAAATCCCTCGAGAATCTTGGACTAGGTTAAACTAGGACAGGATCT 1068

Qy 213 ThreonineAsnThrProLysGluLeu 220

Db 1069 TTGGGGAAACATGGATGGAAACATGGAAATGGAAACCACGAAAGTAGCAATAAAACACTA 1128

Qy 221 AspSerSerLeuLeuHeSerSerLysAlaThrGlyGluLysSerGluGly 240

Db 1129 AAACCCAGGTAAATGATGCGCAGAAGTTCTTCAGAAGCTCAGATAATGAAAAAAATA 1188

Qy 241 LeuArgGluUserLeu--SerPhethyrylIleSerLeuAsnAspGluAlaVal 256

Db 1189 AGACATGATTAACCTTCCACTATATGCTGTTCTGAGAACCAATT 1239

Qy 73 SerGluValSerGluGlyGluLeuSerLeu 90

Db 604 TCAATCGCTACAGGAAGTAATGTCGCCGAAATTATGCCGAGATTCAGGCCAGATTC 663

Qy 91 SerHisGlyTrpLeutyrGluGlyLeuSerArgGluLysAlaGluGluLeu 108  
 Db 664 ATTTCAGGCGAGAAAGTAATGTCGCCGAAATGTCGCCGAAAGATCTGAAAGATTCTT 723

Qy 109 LeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGly 128

Db 724 TGTAAATCTGGAAATCTACGAGGTATTCTCTAGTAAAGAGACTGAAACACTAAAGGT 783

Qy 129 SerTyrosLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArg--- 146

Db 784 GCTTATTCCTTCTTATTCTT 828

Db 889 GCACAAATGGATACTCGAACAAATTGGTGAACACTACAGAACATGCTGATGGTTA 948 Percent Similarity: 52.34%  
 Best Local Similarity: 36.45%  
 Query Match: 23.17%  
 DB: 2 Gaps: 6  
 US-09-939-853A-75 (1-261) x US-09-006-675-1 (1-1491)

Qy 1 Met Gly Ser Leu Pro Ser Arg ----- Arg Ly Ser Ieu Pro Ser Pro Ser Pro Ser 15  
 Db 1 ATGGCTGATCAGTAAAGGTCAATAACACTGGCAAAGTCGGACCTCCGGAA 60  
 Qy 16 Leu Ser Ser Val Glu Gly Pro ----- Val Thr Met --- Glu Ala Glu 31  
 Db 61 AGAACCCAAACCCATTATGTGAGGACCACATCTAACATGACTAAACCTAA 120  
 Qy 32 Arg Ser ----- Lys Ala Thr Ala Val Ala Leu Cys 41  
 Db 121 AGATCATTAAGCACCCAGAGGGAAAGGCAAGAAAGAGTCCTCTGCCTTGAT 180  
 Qy 42 Ser Phe Pro Ala Gly Gly Pro Ala Gly Leu Ser Leu Arg Gly Pro Leu Thr Ile 61  
 Db 181 GACPATGGAGTCACCTGGGATGTGACMTTGGAAAGGGGACCATCCTGCTAA 240  
 Qy 62 Val Ser Glu Asp Phe Ile Phe Ser Gly Val Leu Ser Glu Val Leu Ser Gly Arg Glu Ty Asn 81  
 Db 241 AAGAAAGACTCAGGGAGTGGGGAAACATGCTTAATTTCAACTGGTGAAGAGCTT 300  
 Qy 82 Ile Pro Ser Val His Val Gly Ile Val ----- Ser His Gly Tri Phe Leu Tyr Glu 97  
 Db 301 GTTCCAGTAACATGTTGCTATTCTGAATCTGAAGATGGTACTTAA 360  
 Qy 98 Gly Leu Ser Arg Glu Ile Sal Agl Glu Leu Leu Pro Gly Asn Pro Gly Gly Ala 117  
 Db 361 GGATGAGCGGAGGAGCTGAAGCATGAGCATGAAAGCTTCTCCCTCTGTGCGA----- 420  
 Qy 118 Phenyl Leu Arg Glu Ser Glu Thr Arg Arg Gly Ser Ile Ser Leu Val Arg Leu Ser 137  
 Db 421 TTCTATGATTCGAGACATGAGCATGAAAGCTTCTCCCTCTGTGCGA----- 474  
 Qy 138 Arg Pro Ala Ser Ile Ser Pro Asp Arg Ile Arg His Cys Ile Asp Asn Gly Tyr P 157  
 Db 475 -----GACTCAGGGAACTGTGAAACATTACAAATTGCAACTCGATGAGGGT 528  
 Qy 158 Leu Tyr Ile Ser Pro Leu Thr Phe Pro Ser Leu Glu Ala Leu Val Asp His Tyr Ser 177  
 Db 529 TTCTATGATTCGAGCTCCATACGGATCCTTCTTCAGCTGGTACAGCCATTATCAA 588  
 Qy 178 Glu Leu Ala Asp Asp Phe Ile Cys Cys Leu Leu Ile Asp Glu Pro Cys 191  
 Db 589 GGTAAAGGGATGGCTGTGCTGTGCTTACATAACATGCG 630

RESULT 9  
 US-09-006-675-1  
 Sequence 1, Application US/09006675  
 ; Patents No. 5952213  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hemmati-Birvanli, Ali  
 ; TITLE OF INVENTION: A NOVEL SRC-FAMILY KINASE AND METHODS OF  
 ; TITLE OF INVENTION: USE THEREOF  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; STREET: Klauber & Jackson  
 ; CITY: Hackensack Avenue, 4th Floor  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07601  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,675  
 ; FILING DATE: 13-JUN-1998  
 ; CLASSIFICATION:  
 ; NAME: Jackson Esg, David A.  
 ; REGISTRATION NUMBER: 26,742  
 ; REFERENCE/DOCKET NUMBER: 6000-1-217  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 201-487-5800  
 ; TELEX: 133521  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1491 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: CDNA  
 ; HYPOTHETICAL: NO  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..1491  
 ; Pred. No.: 6.12e-24 Length: 1491  
 ; US-09-006-675-1 Matches: 313.50 Score: 78

RESULT 10  
 US-09-228-603A-1  
 ; Sequence 1, Application US/09228603A  
 ; Patent No. 6291651  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hemmati-Birvanli, Ali  
 ; APPLICANT: Weinstein, Daniel C.  
 ; TITLE OF INVENTION: A NOVEL SRC-FAMILY KINASE AND METHODS OF  
 ; TITLE OF INVENTION: USE THEREOF  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Klauber & Jackson  
 ; STREET: 411 Hackensack Avenue, 4th Floor  
 ; CITY: Hackensack  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07601  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30



Qy 38 -----ValAlaLeuGlySerPheProAlaGlyGlyProAlaGlu 50  
Db 373 GGGATTCGGCTGACCTGTGTTGCGCTGTTGAGCTGAACTTGGAC 432  
Qy 51 LeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu---AspGlyAspTrpTrp 69  
Db 433 CTACCTTCACAAAGGAGAAACTTCACATCTGAACATACTGAGGTACTGTGG 492  
Qy 70 ThrValLeuSerGluValSerGlyArgGluTyroAsnIleProSerValHisValGlyLys 89  
Db 493 GAGGTCTGGCTCTAGCTGGCATCCAGCAACTACCTGGCCCT 552  
Qy 90 Val-----SerHisGlyTyrTrpLeutYrGluGlyLeuSerArgGluLysAlaGlu 105  
Db 553 GTTGACTCAATCCAAAGCTGAAGCTGAGACTGGTACTCTGGAGAG 612  
Qy 106 GluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuGlnThr 125  
Db 613 AGGCAGCTGGCTTTCACCCGGAAACCCCCAGGGCTTCTCATTCGGAAAGGGAGCC 672  
Qy 126 ArgArgGlySerTySerLeuSerValArgLeuSerArgProLaserTrpaspGile 145  
Db 673 ACCAAAGGCGCTACTCCCTGTCATCCGGACTGGGATCACATGTG 732  
Qy 146 ArgHistTyArgLeuHisCysIleLeuAspArgLysTrpLeutYrIleSerProArgLeuThr 165  
Db 733 AAGCTTACAGATCCGAAACTGGACATGGGACTACTACATCACAGGGTCAG 792  
Qy 166 PheProSerLeuGinAlaLeuValAspHisTySerGluLeuAlaASPAspIleCysCys 185  
Db 793 TTCAACTCTGGCTGGAGCTGAGCTAACATGGGGTGAATGAGCTGGTGTGCAAC 852  
Qy 186 LeuLeuLysGluProCysValLeuLeuGlnArg 195  
Db 853 CTGCTCATGGCCCTGCACTATGAG 882

RESULT 12  
US-09-470-881-2 Application US/09470881  
; GENERAL INFORMATION:  
; Parent No. 6/855328  
; APPLICANT: ELICEIRI, Brian  
; APPLICANT: CHREESH, David A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR MODULATION OF  
; TITLE OF INVENTION: ANGIOGENESIS AND VASCULAR PERMEABILITY USING SRC OR  
; FILE REFERENCE: TSP1 651.2  
; CURRENT FILING DATE: 1999-12-22  
; PRIOR APPLICATION NUMBER: PCT/US99/11780  
; PRIOR FILING DATE: 1999-05-28  
; PRIOR APPLICATION NUMBER: 60/087,220  
; PRIORITY FILING DATE: 1998-05-29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO. 2  
; LENGTH: 1759  
; TYPE: DNA  
; ORGANISM: Chicken  
; FEATURE:  
; NAME/KEY: Gene  
; LOCATION: (1) : (1759)  
; OTHER INFORMATION: chicken c-SRC cDNA  
; NAME/KEY: CDS  
; LOCATION: (112)..(1710)  
; US-09-470-881-2

Alignment Scores:  
Pre. No.: 3.35e-21  
Score: 289.00  
Percent Similarity: 47.33%  
Best Local Similarity: 30.60%  
Query Match: 21.36%

Gaps: 4  
Length: 1759  
Matches: 86  
Conservative: 47  
Mismatches: 104  
Indels: 45

Db: US-09-939-853A-75 (1-261) x US-09-470-881-2 (1-1759)  
Qy 5 ProSerArgArgLysSerIleProSerProSerIleSerSerValGlnGlyGlnGly 24  
Db 317 CGTTPAGTCGCCGGCGACTGGCTGGCGCGTCAACC----- 363  
Qy 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44  
Db 364 -----ACTTCGTCGGCTCTAGACTACGACTACGAG 390  
Qy 45 AlaGlyGlyProAlaGluUserLeuArgLeuGlyGluProLeuThrIleValSerGlu 64  
Db 391 TCCCGACTGAAACGACATGGTCCTTAAGAAAGGAAAGCCGCTGAGATTGCAACAC 450  
Qy 65 ---AspGlyAspTrpTrpTrpValLeuSerGluValSerIleArgGluTyrasNlePro 83  
Db 451 ACGGAAAGGIGACTGTGGCTCATCCCTACTACAGGACAGGCGCTAAATCCCC 510  
Qy 84 SerValHisValGlyLys-----ValSerHisGlyTrpLeutYrGluGlyLeu 99  
Db 511 AGTAACATGTCGCCCTAGACTTCATTCAGCTGAAGAGTGACTCTGGAAAGATC 570  
Qy 100 SerArgGlyLysAlaGluGluLeuLeuProGlyAsnProGlyLysAlaPheSer 119  
Db 571 ACTCCTGGCAAGTGGCTGGAGCTGAGGAGCTGCTGCTAACCCGAAACCCGGAAACCTTCCTG 630  
Qy 120 IleArgGluSerGlnThrArgArgGlySerTySerIleSerValArgLeuSerArgPTO 139  
Db 631 GTCCTGGAGAGGAGACGACAAAGTGCCTATTCGCTTCCTGGCTTTCGACTTGTGACAC 690  
Qy 140 AlaSerTrpAspArgLysLeuGlyLeuGlyLeuGlyLeuGlyLeuGlyLeu 159  
Db 691 GCCAAGGGCTCAATGGAACTAGATCGCANGCTGGCAGGGCTTCAPAC 750  
Qy 160 IleSerProArgLeutYrIleLeuGlyLeuGlyLeuGlyLeuGlyLeu 179  
Db 751 ATCACCTCAGGCCACAGCTAGAGCTTCAGCAGCTGGCAGCTACTCCAAAAT 810  
Qy 180 AlaAspAspIleCysCysLeuLeuLysGluProCys----- 191  
Db 811 GCTGATGGCTGTGCAAGCCCTGACCAAACGTCGAAACGCTGCCCACAGCCAG 870  
Qy 192 ---ValLeuGlnArgAlaGlyProLeuProLeuProLeuProValThrVal 210  
Db 871 GGACCTGGCAGAGGAGCTGGGAAATCCCGGGAGTGTCTGGCTGGCTGGCTGG 930  
Qy 211 GlnArg-----ThrProLeuAsn 216  
Db 931 CGGCCAGGGCTGCTGGAGGGCTGGATGGGACCTGGACACGAGTCGGCC 990  
Qy 217 TrpLyGluLeuAspSerIleLeuPheserGluAlaAlaThrGlyGluGluSerLeu 236  
Db 991 ATAAGAGACTCTGAAGGCCGACCATGTCCTGGAGGCTTCCTGCAAGAAGCCCACTG 1050  
Qy 237 LeuSerGluGlyLeuArgGluUserLeu--SerIlePheLeuLeuAspGluAla 255  
Db 1051 ATGAAAGAACGCTCCGGCATGAGAAGCTGGTCAGCTGAGCTGTTGAGAGGCC 1110  
Qy 256 Val 256  
Db 1111 ATC 1113

RESULT 13  
US-07-820-011A-1  
; Sequence 1, Application US/07820011A  
; Patent No. 5336615  
; GENERAL INFORMATION:  
; APPLICANT: Bell, Leonard  
; APPLICANT: Madri, Joseph A.  
; APPLICANT: Warren, Stephen L.  
; APPLICANT: Luthringier, Daniel J.

TITLE OF INVENTION: Genetically Engineered  
Endothelial Cells Exhibiting Enhanced  
Migration

TITLE OF INVENTION: and Plasminogen Activator Activity

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:  
STREET: 1951 Burr Street  
CITY: Fairfield  
COUNTRY: Connecticut  
ZIP: 06430

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb storage  
COMPUTER: IBM PC XT  
OPERATING SYSTEM: PC-DOS/MS-DOS 2.10  
SOFTWARE: Displaywrite 3

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,011A  
FILING DATE: 19/20/06  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: LB-101

TELECOMMUNICATION INFORMATION:  
TELEFAX: (203) 255 1400

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1602 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
ANTI-SENSE: No  
HYPOTHETICAL: No

ORIGINAL SOURCE:  
ORGANISM: Gallus, gallus

PUBLICATION INFORMATION:  
AUTHORS: Takeya, Tatsuo  
TITLE: Structure and Sequence of the  
Cellular Gene Homologous to the RSV src  
TITLE: Gene and the Mechanism For Generating the  
Transforming Virus

JOURNAL: Cell  
VOLUME: 32  
PAGES: 881-890  
DATE: March, 1983  
US-07-820-011A-1

Alignment Scores:  
pred. No.: 5 ProSerArgArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24  
Score: 1602  
Percent Similarity: 51.28%  
Best Local Similarity: 34.62%  
Query Match: 21.21%  
DB: 1

US-09-939-853a-75 (1-261) x US-07-820-011A-1 (1-1602)

Length: 1602  
Matches: 81  
Conservative: 39  
Mismatches: 84  
Index: 32  
Gaps: 5

Qy 5 ProSerArgArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24  
Db 206 CCGTTAGTCGGCCGCGAGCTGGGACTGGCTGGTCAACCTGGTACCC----- 252

Qy 25 ProValThrMetGluAlaGluArgSerLysAlaLysGluSerPhePro 44  
Db 253 -----ACTTGCTGGCTCTCTAGACTACBAG 279

Qy 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrLeuSerLeu 64  
Db 280 TCCCCGAACTGAACGGGACTGCTCTTCAAAAGGAGAACGCCCTGAGATGTGAAACAC 339

Qy 65 ---AspGlyAspartPtpThrValLeuSerGluValSerGlyArgGluTyrosAsnIlePro 83  
Db 340 ACGGAAGGTCGACTGTCGCTGGCTCATCTCTACTAGAACGACGGCTACATCCC 399

Qy 84 SerValHisValGlyLys-----ValserHisGlyLysTrpLeutyrGluGlyLeu 99  
Db 400 AGTAACATATCGGCCCTCAGACTCCATCAGCTGAGAGCTGTTGAAAGATC 459

Qy 100 SerArgGluSerAlaGluGluLeuLeuProGlyAlaPheLeu 119  
Db 460 ACTCGTGGAGGAGTCGGCTGAGCGCTCTGCTAACCCGAAACCCGGAAACCTCTTG 519

Qy 120 IleArgGluSerGlnThrArgArgGlySerTySerLeuSerValArgLeuSerArgPro 139  
Db 520 GTCGGGAGAGGAGACGAAACCCGAAACCCGAAACCCGGAAACCTCTTGACTTGACAC 579

Qy 140 AlaSerTerPASPArgSerGlySerLysLeuAspArgGlyLysTrpLeutyr 159  
Db 580 GCCAAAGGGCTCAATGCAAGCTGAACTGGCAAGCTGGCTTCTACATCCAAACAT 639

Qy 160 IleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTrpSerGlnLeu 179  
Db 640 ATCACCTAGGCACAGCTCAGGCTCAGCAGCTGGCTACTACTCCAAACAT 699

Qy 180 AlaAspAspLysCysSerLeuLeuLysGluProCys----- 191  
Db 700 GCTGATGGCTGGTGGCTGACGTCGGCCACGTCAGCTGGCTACTACTCCAAACAT 759

Qy 192 ---ValLeuGlnArgAlaGlyProLeuProGlyLysAspPheProLeuProValThrVal 210  
Db 760 GGACTCGGCCAGGAGCGCTGGGAAATCCCGGGAGTCGCTGGCTGAGGCTGAGGCTG 819

Qy 211 GluAspThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224  
Db 820 GGGCAGGGCTG-CIT-----TGAGAGGCTCTGGATGGGACCTG 857

RESULT 14  
PCT-US93-00445-1  
; Sequence 1, Application PC/TUS9300445

GENERAL INFORMATION:  
APPLICANT: Bell, Leonard  
APPLICANT: Madri, Joseph A.  
APPLICANT: Warren, Stephen L.  
APPLICANT: Lutherling, Daniel J.  
TITLE OF INVENTION: Genetically Engineered  
TITLE OF INVENTION: Endothelial Cells  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 760 Kb storage  
COMPUTER: DELL 486/50  
OPERATING SYSTEM: DOS 5.0  
SOFTWARE: Displaywrite 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00445  
FILING DATE: 1993/01/05  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00445  
FILING DATE: 07/820,011  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-101PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255 1400

TELEFAX: (203) 254 1101  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1602 base pairs  
 TYPE: NUCLEAR ACID  
 STRANDEDNESS: Double  
 TOPOLOGY: Linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: No  
 ANTI SENSE: No  
 ORIGINAL SOURCE:  
 ORGANISM: Gallus gallus  
 PUBLICATION INFORMATION:  
 AUTHORS: Takeya, Tatsuo  
 AUTHOR: Hanafusa, Hidesaburo  
 TITLE: Structure and Sequence of the Cellular Gene Homologous to the RSV src Gene and the Mechanism for Generating the Transforming Virus  
 JOURNAL: Cell  
 VOLUME: 32  
 PAGES: 881-890  
 DATE: March, 1983  
 PCT-US93-00445-1

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
QY 5	4.75e-21	1602	81	39	84	32	5
Db 206	287.00	81	39	34	62	21	18
QY 25	51.28%	39	34	62	21	21	8
DB: 253	Best Local Similarity:	34.62%	34	62	21	21	8
DB: 253	Query Match:	34.62%	34	62	21	21	8

US-09-939-853a-75 (1-261) × PCT-US93-00445-1 (1-1602)

QY 5 ProserArgArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGlyGly 24  
 Db 206 CCGTGTACGGCAGCTGGGGGGACTGTGGCCCTGACCTGGGGCTCACCC----- 252  
 QY 25 ProValThrMetGluValGluLysSerLysAlaThrAlaAlaLeuGlySerPhePro 44  
 Db 253 -----ACTTCGTTGGCTCTAGACTACAGAG 279  
 QY 45 AlaGlyGlyProAlaGluLeuSerLeuProSerProSerLeuSerThrIleValSerGlu 64  
 Db 280 TCCGGGACTGAAAGGACTTGCTCTTAAGAAGAGGAAACGCTGGAGATTCACAAAC 339  
 QY 65 --AspGlyAspProTrpThrValLeuSerGluValSerGlyArgLysTyrAsnIlePro 83  
 Db 340 ACCGAAGGTGACTGTGGCTGGCTCATCCCTCACTACAGAACAGCGGGTACATCCCC 399  
 QY 84 SerValHisValGlyLys-----ValSerHisGlyTryptophylGluGlyLeu 99  
 Db 400 AGTAACTATGTCGGGCCCTCAGCTCATCAGCTGAAAGATGGTACTTGGGAGATC 459  
 QY 100 SerArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAlaProGlyAlaPheLeu 119  
 Db 460 ACGTCGGAGTCGAGCTGACGGCTCTGCTCAACCCGAAAACCCGGGGAAACCTCTTG 519  
 QY 120 IleArgGluSerGlnThrArgGlySerIleSerLeuSerValArgLeuSerArgPro 139  
 Db 520 GTCCGGGGAGGGAGGACCAAAAGGTGCCATTGCTCTCGTCTGACTTTGACAAC 579  
 QY 140 AlSerTrpAspArgIleArgHsTyrArgIleIleSlysLeuAspAsnGlyTryptLeuTyR 159  
 Db 580 GCGAAAGGGGCTCAATGCAAGATCGCCAGGACAGCTGGCTCTCTAC 639  
 QY 160 IleSerProArgLeuThrPheProSerLeuLeuValAspIleSlysGlySerGluLeu 179  
 Db 640 ATCACCTTACGCCACAGTTGAGCTGGCTGAGCTGGCTACTACTCCAAACAT 699  
 QY 180 AlaAspAspIleCysCysLeuIleLysGluLeuProCys----- 191

RESULT 15  
 US-07-820-011A-3  
 / Sequence 3, Application: US/07B20011A  
 / GENERAL INFORMATION:  
 / APPLICANT: Bell, Leonard  
 / APPLICANT: Madri, Joseph A.  
 / APPLICANT: Warren, Stephen L.  
 / APPLICANT: Luthringer, Daniel J.  
 / TITLE OF INVENTION: Genetically Engineered Endothelial Cells Exhibiting Enhanced Migration  
 / NUMBER OF SEQUENCES: 4  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Maurice M. Klee  
 / STREET: 1951 Burr Street  
 / CITY: Fairfield  
 / STATE: Connecticut  
 / COUNTRY: USA  
 / ZIP: 06430  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: 5.25 inch, 360 Kb storage  
 / COMPUTER: IBM PC XT  
 / OPERATING SYSTEM: PC-DOS/MS-DOS 2.10  
 / SOFTWARE: Displaywrite 3  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/07/820,011A  
 / FILING DATE: 19920106  
 / CLASIFICATION: 435  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Maurice M. Klee  
 / REGISTRATION NUMBER: 30,399  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (203) 255 1400  
 / TELEFAX: (203) 254 1101  
 / INFORMATION FOR SEQ ID NO: 3:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 1611  
 / TYPE: NUCLEIC ACID  
 / STRANDEDNESS: Double  
 / TOPOLOGY: Linear  
 / MOLECULE TYPE: cDNA to mRNA  
 / HYPOTHETICAL: No  
 / ANTI-SENSE: No  
 / ORIGINAL SOURCE: Homo sapien  
 / POSITION IN GENOME:  
 / CHROMOSOME/SEGMENT: Chromosome 20  
 / PUBLICATION INFORMATION:  
 / AUTHORS: Anderson, Stephen K.  
 / AUTHORS: Gibbs, Carol P.  
 / AUTHORS: Tanaka, Akio  
 / AUTHORS: Kung, Hsing-Jien  
 / AUTHORS: Fujita, Donald J.  
 / TITLE: Human Cellular Src Gene:  
 / TITLE: Nucleotide Sequence and Derived Amino Acid Sequence of the Region Coding for the Carboxy-Terminal Two-Thirds of  
 / TITLE: PPP0-C-SRC  
 / JOURNAL: Molecular and Cellular Biology  
 / VOLUME: 5

ISSUE: 5  
 PAGES: 1122-1129  
 DATE: May, 1985  
 AUTHORS: Tanaka, Akio  
 Gibbs, Carol P.  
 AUTHORS: Arthur, Richard R.  
 AUTHORS: Anderson, Stephen K.  
 AUTHORS: Kung, Hsing-Jien  
 AUTHORS: Fujita, Donald J.  
 TITLE: DNA Sequence Encoding the Amino-Terminal Region of the Human C-src Protein: Implications of Sequence Divergence among src-Type Kinase Oncogenes  
 JOURNAL: Molecular and Cellular Biology  
 VOLUME: 7  
 ISSUE: 5  
 PAGES: 1978-1983  
 DATE: May, 1987  
 US-07-820-011A-3

Alignment Scores:

Qy	8	ArgLySerIleProSerProSerLeuSerSerSerValGlyGlyGlyProValThr	27
Score:	6.41e-20	Length:	16:1
Percent Similarity:	276.50	Matches:	87
Best Local Similarity:	46.15%	Conservative:	45
Query Match:	30.42%	Mismatches:	117
DB:	20.44%	Indels:	38
	1	Gaps:	6

US-09-939-853A-75 (1-261) x US-07-820-011A-3 (1-1611)

Qy	8	ArgLySerIleProSerProSerLeuSerSerSerValGlyGlyGlyProValThr	27
Score:	6.41e-20	Length:	16:1
Percent Similarity:	276.50	Matches:	87
Best Local Similarity:	46.15%	Conservative:	45
Query Match:	30.42%	Mismatches:	117
DB:	20.44%	Indels:	38
	1	Gaps:	6

US-09-939-853A-75 (1-261) x US-07-820-011A-3 (1-1611)

Qy	8	ArgLySerIleProSerProSerLeuSerSerSerValGlyGlyGlyProValThr	27
Score:	6.41e-20	Length:	16:1
Percent Similarity:	276.50	Matches:	87
Best Local Similarity:	46.15%	Conservative:	45
Query Match:	30.42%	Mismatches:	117
DB:	20.44%	Indels:	38
	1	Gaps:	6

Search completed: February 20, 2004, 11:33:00  
 Job time : 120 secs

---

754 AAGCCCGAGACTCAGGCCCTGGAGATCCCTGGAGTCGCTGGCGG 813  
 206 LeuProValThrValGlnArg-----212  
 814 CTGGAGTCAGCTGGCCAGGCTGATGGCTGCTGGATGGAAACGGT 873  
 213 ---ThrProLeuAsnTrpIlysGluLeuAspSerLeuLeuPheSerGluAlaAlaThr 231  
 874 ACCACAGGGGGCCATCAAACCTGAACCTGCAACATGCTGACAGGCTGCTCG 933  
 232 GlyGluGluSerLeuSerGluGlyLeuArgGluSerLeu--SerPheTyrIleSer 250  
 934 CAGGAGCCCAGGTGAGGATGAGCTGAGCATGAGAAGCTGGTGTATGCTGTC 993  
 251 LeuAsnAspGluAlaVal 256  
 994 GTTTCAAGGGAGCCCCATT 1011  
 Db 994 GTTTCAAGGGAGCCCCATT 1011

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 10:48:38 ; Search time 357 Seconds  
(Without alignment!)  
2560.026 Million cell updates/sec

Title: US-09-939-853a-75  
Perfect score: 1353  
Sequence: 1 MGSLSPRRKSLPSLSSSV.....RESLSFYISINDEAVSLDDA 261

Scoring table: BIOSUM62  
Xgapext 10.0 , Xgapext 0.5  
Ygapext 10.0 , Ygapext 0.5  
Fgapext 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2308684 seqs, 175022206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+pn.model -DEV=xlp  
-Q=/con2\_1/USPTO\_spool/p/US0939853@runat\_19022004\_145341\_24590/app\_query.fasta\_1.455  
-DB=Published\_Applications\_NA -QMMR=fastrap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCH=0 -LOOPTXT=0 -UNITS=bits  
-TRANS=human40\_cdi -LIST45 -DOCALIGN=200 -THR SCORE=pot -THR MAX=100  
-THR MIN=0 ALIGN=15 MODE=LOCAL\_OUTEMT=pto NORM=ext -HEPST2E=500 -MINLEN=0  
-MAXLEN=200 -NC\_MMAP=19022004\_145341\_24590  
-NCPU=6 -ICPU=3 -NC\_MMAP=19022004\_145341\_24590  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELEXIT=7

Database : Published\_Applications\_NA: \*

```

1: /cgn2_6/pctodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/pctodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/pctodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/pctodata/2/pubpna/US05_PUBCOMB.seq:*
5: /cgn2_6/pctodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/pctodata/2/pubpna/US08_PUBCOMB.seq:*
7: /cgn2_6/pctodata/2/pubpna/US09_NEW_PUB.seq:*
8: /cgn2_6/pctodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/pctodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/pctodata/2/pubpna/US08_PUBCOMB.seq:*
11: /cgn2_6/pctodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/pctodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/pctodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/pctodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/pctodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/pctodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/pctodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/pctodata/2/pubpna/US60_PUBCOMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	99.6	786	14	US-10-043-649-1
2	61.0	763	9	US-09-967-550-953
3	47.5	864	10	US-09-814-53-1202
4	58.6	875	9	US-09-867-550-1915
5	48.8	3756	13	US-10-002-000-91
6	48.7	2665	9	US-09-854-456-499
7	45.2	444	9	US-09-867-551-951
8	370.5	2298	14	US-10-175-123-50
9	370.5	27.4	15	US-10-159-563-343
10	360.5	26.6	15	US-10-193-720-1
11	360.5	26.6	2015	14 US-10-007-010-0-3
12	360.5	26.6	2341	14 US-10-25-157-140
13	360.5	26.6	2129	10 US-09-960-706-954
14	358.5	26.5	1911	9 US-09-917-800A-1611
15	358.5	26.5	150	10 US-09-914-353-1731
16	350	25.9	320	10 US-09-914-353-1731
17	348	2032	17	US-10-16-388-27
18	340	25.1	2017	15 US-10-062-674-1776
19	340	25.1	2129	10 US-09-960-706-954
20	340	25.1	2129	15 US-10-305-152
21	324	23.9	2333	14 US-10-24-965-114
22	322.5	23.8	2451	9 US-09-771-161A-4
23	320	23.7	1609	9 US-09-771-161A-30
24	320	23.7	1995	9 US-09-771-161A-31
25	320	23.7	4414	14 US-10-101-510-512
26	315.5	23.3	4343	15 US-10-159-563-184
27	315.5	23.3	4517	14 US-10-198-377A-3
28	314.5	23.3	4517	14 US-10-007-266A-104
29	314.5	23.2	549	10 US-09-918-995-25044
30	312.5	23.1	1590	15 US-10-085-117-18
31	312.5	23.1	2354	9 US-09-967-761-2829
32	312.5	23.1	2354	15 US-10-153-650-123
33	312.5	23.1	2354	15 US-10-085-117-17
34	303	22.4	1554	15 US-10-085-117-15
35	303	22.4	2179	15 US-10-085-117-14
36	283.5	21.0	488	10 US-09-918-995-19745
37	246	18.2	432	9 US-09-884-761-2829
38	246	18.2	448	9 US-09-884-761-2829
39	240.5	17.8	2770	9 US-09-977-269-5
40	240.5	17.8	2770	9 US-19-977-269-5
41	240.5	17.8	2770	10 US-09-977-269-5
42	240.5	17.8	2863	9 US-09-984-455-1631
c	240.5	17.8	7607	10 US-09-982-610-19
44	238.5	17.6	2173	15 US-10-094-749-1047
45	230.5	17.0	342	15 US-10-062-574-1242

Sequence 1, APP1  
Sequence 953, APP1  
Sequence 1913, APP1  
Sequence 1915, APP1  
Sequence 91, APP1  
Sequence 499, APP1  
Sequence 951, APP1  
Sequence 50, APP1  
Sequence 343, APP1  
Sequence 1, APP1  
Sequence 1983, APP1  
Sequence 140, APP1  
Sequence 3, APP1  
Sequence 1611, APP1  
Sequence 17314, APP1  
Sequence 27, APP1  
Sequence 31, APP1  
Sequence 512, APP1  
Sequence 184, APP1  
Sequence 3, APP1  
Sequence 4, APP1  
Sequence 30, APP1  
Sequence 31, APP1  
Sequence 1776, APP1  
Sequence 1776, APP1  
Sequence 954, APP1  
Sequence 1452, APP1  
Sequence 114, APP1  
Sequence 17, APP1  
Sequence 15, APP1  
Sequence 25044, APP1  
Sequence 18, APP1  
Sequence 300, APP1  
Sequence 123, APP1  
Sequence 17, APP1  
Sequence 15, APP1  
Sequence 14, APP1  
Sequence 19745, APP1  
Sequence 19745, APP1  
Sequence 2829, APP1  
Sequence 15513, APP1  
Sequence 5, APP1  
Sequence 1631, APP1  
Sequence 19, APP1  
Sequence 1087, APP1  
Sequence 1242, APP1

## ALIGNMENTS

RESULT 1  
US-10-043-649-1  
; Sequence 1, Application US/10043649  
; Publication No. US2003005992AA1  
; GENERAL INFORMATION:  
; APPLICANT: Holland, Sacha J.  
; APPLICANT: Mendenhall, Marcy K.  
; APPLICANT: Pardo, Jorge  
; APPLICANT: Spencer, Collin  
; APPLICANT: Sheng, Ning  
; APPLICANT: Liao, X. Charlene  
; APPLICANT: Wu, Jun  
; APPLICANT: Zhou, Xiulan  
; APPLICANT: Shen, Mary  
; APPLICANT: Liao, X. Charlene  
; TITLE OF INVENTION: Cloning of a No. US20030059924A1 Inhibitor of Antigen-receptor Screen  
; FILE REFERENCE: A-70219-1/RMS/DHR  
; CURRENT APPLICATION NUMBER: US/10/043, 649  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/260, 953



RESULT 3  
US-09-814-353-21302      Sequence 21302, Application US/09814353  
Publication No. US20030165831A1      GENERAL INFORMATION:  
APPLICANT: Lee, John      APPLICANT: Thompson, Pamela  
APPLICANT: Lillie, James      APPLICANT: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
FILE REFERENCE: MRI-006B  
CURRENT APPLICATION NUMBER: US/09/814,353  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/191,031  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: US 60/207,124  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: US 60/211,940  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 60/216,820  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: US 60/220,661  
PRIOR FILING DATE: 2000-07-25  
PRIOR APPLICATION NUMBER: US 60/257,672  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 22037  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO: 21302  
LENGTH: 864  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME: misc\_feature  
LOCATION: 1, 2-, 3, 32, 862, 863, 864  
OTHER INFORMATION: n = A,T,C OR G  
US-09-814-353-21302

Alignment Scores:  
Pred. No.: 1. 54e-64      Length: 864  
Score: 643.00      Matches: 126  
Percent Similarity: 98.45%      Conservative: 1  
Best Local Similarity: 97.67%  
Query Match: 47.52%  
DB: 10  
Gaps: 0

US-09-939-853A-75 (1-261) × US-09-814-353-21302 (1-864)

Qy 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerSerVal 20  
Dbb 4550 ATGCCAAGCTGCCCCATGCAAGAAATCTCTGCCAACGCCAAGTTGGATTCCCTCTGTC 509  
Qy 21 GlnglyGlnGlyProValThrMetGluAlaGluArgSerLysAlaIvalAlaLeu 40  
Db 510 CAAGCCAGGACCTGTGACATGAGGAGGACACGCCGAGGCCG 569  
Qy .4 GlySerPheProAlaGlyGlyProAlaGlyLeuSerLeuArgLeuGlyLeuProLeuthr 60  
Dbb 570 GGCACTTCCGGAGGTGGCCGCCGGCTGCTGAGACTGGGAAGCCATTGAC 629  
Qy 61 IleValSerGluAsnGlyAspTrpTrpThrValLeuSerIleValSerIleArgGlyLeu 80  
Dbb 630 ATCGCTCTTGGATTGAGGAGAAGCTGGAACTGGCTGAGCTCAGGCAGAGAT 689  
Qy 81 AsnIleProSerValHisValGlyLysValSerIleSerIleGlyLeuSer 100  
Dbb 690 AACATCCCAGGNCACGGCCAAAGCTTCCCATGGGGCTGTTAGCTGAC 749  
Qy 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaphenile 120  
Dbb 750 AGGGAAAAGCCAGGAAACTGGAAACCTGGAACTGGCTGTTACCTGGAACTGGCC 809  
Qy 121 ArgGluSerGlnThrArgArgGlySer 129  
Db 810 CGGAGAGCCAGACCCAGGAGAGAGCTGC 836

RESULT 4  
US-09-867-550-1915      Sequence 1915, Application US/09867550  
Patent No. US20030082206A1      GENERAL INFORMATION:  
APPLICANT: Leach, Martin D.      APPLICANT: Mehraban, Fuad.  
APPLICANT: Conley, Pamela      APPLICANT: Law, Debbie  
APPLICANT: Topper, James      APPLICANT: Cura-313  
TITLE OF INVENTION: Theraby      CURRENT APPLICATION NUMBER: US/09/887,550  
FILE REFERENCE: 214402-013 (Cura-313)  
CURRENT FILING DATE: 2001-09-20  
PRIOR APPLICATION NUMBER: USSN 60/208,427  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 2125  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO: 1915  
LENGTH: 875  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)  
OTHER INFORMATION: Wherein n is one of a or t or c or g  
US-09-867-550-1915

Alignment Scores:  
Pred. No.: 6.02e-58      Length: 875  
Score: 586.00      Matches: 112  
Percent Similarity: 100.00%      Conservative: 0  
Best Local Similarity: 100.00%  
Query Match: 43.31%  
DB: 9  
Gaps: 0  
Indels: 0

US-09-939-853A-75 (1-261) × US-09-867-550-1915 (1-875)

Qy 150 IleHsCysIleAspAsnGlyTryptophylLeuProArgLeuThrPheProSerLeu 169  
Db 4 ATCCPCTGCTTGACATGGCTGGTGACATCTACCGGCCAACCTCCACTC 63

Qy 170 GluAlaLeuValAspHistyrSerGluLeuAlaAspAspIleCysCysIleLeuIysGlu 189  
Db 64 CAGGCCCTGCTGGACATTAATCTGAGCTGGCTACTCTAGGAG 123  
Qy 190 ProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThr 209  
Db 124 CCCTCTGTCTGCAGAGGGCTGCCGCCCTCCCTCAAGCATCACCCCTACCTCTGACT 183  
Qy 210 ValGlnArgThrProLeuAsnTryptophylLeuAspSerLeuLeuIleSerIleLeu 229  
Db 184 GTCGAGGACACCACTGAAAGGGCTGAACTGAAAGGGCTGAAAGCTGTTTCGAGCT 243  
Qy 230 AlaThrGlyGluGluSerLeuSerGluGlyLeuArgGluSerLeuSerPheTyryle 249  
Db 244 GCCAGGGGGAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTACATC 303  
Qy 250 SerLeuAspAspGluIalaValSerLeuAspAspAla 261  
Db 304 AGCTGAATGAGGAGCTCTCTCTGATGATGATC 339

RESULT 5  
US-10-000-600-91      Sequence 91, Application US/10002600  
Publication No. US20020137077A1      GENERAL INFORMATION:  
APPLICANT: Hopkins, Christopher M.      APPLICANT: Peterson, David P.  
APPLICANT: Perkins, Christopher M.

APPLICANT: Cocks, Benjamin G.  
 APPLICANT: Hawkins, Phillip R.  
 TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS  
 FILE REFERENCE: PA-0042 US  
 CURRENT APPLICATION NUMBER: US/10/002,600  
 CURRENT FILING DATE: 2001-10-25  
 PRIOR APPLICATION NUMBER: 60/243,521  
 PRIOR FILING DATE: 2000-10-25  
 NUMBER OF SEQ ID NOS: 116  
 SOFTWARE: PERL Program  
 SEQ ID NO: 91  
 LENGTH: 3756

FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Template ID: 059263.15  
 US-10-032-600-91

ORGANISM: Homo sapiens

Alignment Scores:

Prd. No.:	9.54e-46	Length:	3756
Score:	488.00	Matches:	101
Percent Similarity:	57.20%	Conservative:	46
Best Local Similarity:	39.30%	Mismatches:	94
Query Match:	36.07%	Indels:	16
DB:	13	Gaps:	3

US-09-939-853a-75 (1-261) × US-10-002-600-91 (1-3756)

Qy 5 ProSerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24  
 Db 1098 CCGGGAAAAGAAAATGGAAAACAGCATAACTCCATGCCGAGGG 1157

Qy 25 ProValThrMetGluAlaGluArgSerIysAlaValAlaLeuGlySerPhePro 44  
 Db 1158 CCCCTGCCAACCGGGACTGGATGGACTGGACTGGCTGCTAAAGGACTACCCG 1217

Qy 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyLeuProLeuThrIleValSerGlu 64  
 Db 1218 TCTCTGACATCAAGCCCCGATATTCCGCGGGGAAACTCGCTGTGTTCTGAT 1277

Qy 65 AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyrosineProSer 84  
 Db 1278 GAACGGGGGGGGAAAACSTARTCTCTAGACTGTCGAGAATTACATCCCTGGA 1337

Qy 85 ValHisValGlyLysValSerHisGlyProLeuTyroGluGlyLeuSerArgGluLysAla 104  
 Db 1338 ATATCGTGCCTCCAGTGTACCTAGCTGCTCTTGCGCTGGCAGACAGGCC 1397

Qy 105 GluGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuLeuArgGluSerGln 124  
 Db 1398 GAGGAGCTGCTGCGAGCTGCCAGACAAGGCGCTCTCATGATCAGAGAGCTGAG 1457

Qy 125 ThrArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144  
 Db 1458 ACCAAARGGTTTACTCACTTGCGTGAAGCAAAGG-  
 CAG 1499-

Qy 145 IleArgHistYrArgIleIleScylsLeuAspAspGlyTyrIleSerProArgLeu 164  
 Db 1500 GTAAGCTTACCGATTTCGGTGTGCCAACACTGTTACATTCCGAGGCTC 1559

Qy 165 ThrPheProSerLeuGlnAlaLeuValAspHistYrSerGluLeuaAspAspIleCys 184  
 Db 1560 ACCTCAGTGGCTGGAGAACCTATTGAGGGCTGTAAGCCCTGTC 1619

Qy 185 CysLeuLeuGlnArgAlaGlyProLeuProGlyIysAspIle 204  
 Db 1620 TGTGTCCTCACACGCCCTGAGAACAGCTGGCTGCCAGCAGTGGCTCC 1679

Qy 205 ProLeuProValThrValGlnArgThrProLeuAsnTrpLeuAspSerSerLeu 224  
 Db 1680 AGCTCACCTTCACCTTGCCCTCAGAGACTGTGAGAGTGTCACTGAGA 1733

Qy 225 LeuPheSerGluAlaAlaThrGly-----GluGluSerLeuLeu 237  
 Db 1734 ---CTGCAAGAGACCCGAGAACAGAGAACCCGCTGAGTCCTTTC 1790

Qy 238 SerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAspGlu 254  
 Db 1791 AGCTATGCCCTCGAGAGCATCTACCTGCTGACCTGAG 1841

RESULT 6  
 US-09-939-456-499  
 Sequence 439, Application US/0954456  
 Patent No. US20020115057A1.

GENERAL INFORMATION:  
 APPLICANT: Young, Paul  
 TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cancer Cells  
 FILE REFERENCE: 689390-76  
 CURRENT APPLICATION NUMBER: US/09/954,456  
 PRIOR APPLICATION NUMBER: US/60/233,617  
 PRIOR FILING DATE: 2000-09-18  
 PRIOR APPLICATION NUMBER: US/60/234,052  
 PRIOR FILING DATE: 2000-09-20  
 PRIOR APPLICATION NUMBER: US/60/234,923  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: US/60/235,134  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: US/60/235,637  
 PRIOR FILING DATE: 2000-09-26  
 PRIOR APPLICATION NUMBER: US/60/235,638  
 PRIOR FILING DATE: 2000-09-26  
 PRIOR APPLICATION NUMBER: US/60/235,711  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235,720  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235,840  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: US/60/235,863  
 PRIOR FILING DATE: 2000-09-27  
 NUMBER OF SEQ ID NOS: 2276  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO: 499  
 LENGTH: 2665  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-954-456-499

Alignment Scores:

Frd. No.:	7.73e-46	Length:	2665
Score:	487.00	Matches:	101
Percent Similarity:	57.20%	Conservative:	46
Best Local Similarity:	39.30%	Mismatches:	94
Query Match:	35.99%	Indels:	16
DB:	9	Gaps:	3

US-09-939-853A-75 (1-261) × US-09-954-456-499 (1-2665)

Qy 5 ProSerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24  
 Db 24 CCAGGGAAAGAAAGAAAGAAACGATGAAATGCCCTGCGCTGGAGAGCG 83

Qy 25 ProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeuGlySerPhePro 44  
 Db 84 CCCCTCCCPAACCCGAGGGACTGGATAGCGACTTCCTGCTAAGTGAATCCCG 143

Qy 45 AlaGlyGlyProAlaGluLeuSerLeuGlyLeuProLeuThrIleValSerGlu 64  
 Db 144 TCTCTGACATCAGGCCCGATATTCCGCGAGGGAAACTGGCTGATTTCTGAT 203

Qy 65 AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyroAsnIleProSer 84  
 Db 204 GAAGGGGGCTGGTGGAAAAGTCTAGACTGGTCACTGTCATCCCTGA 263





Alignment Scores:  
Pred. No.: 2.02e-31 Length: 1924  
Score: 360.50 Matches: 77  
Percent Similarity: 58.38% Conservative: 31  
Best Local Similarity: 41.62% Mismatches: 70  
Query Match: 26.64% Indels: 7  
DB: 15 Gaps: 2

US-09-939-853A-75 (1-1924) x US-10-193-720-1 (1-1924)

Qy 12 ProserProSerLeuSerSerValGlnGlyProvalThrMetGluAlaGlu 31  
Db 195 CGGGGCTATAAGCCNCAAAAGC 245

Qy 32 ArgSerIysValAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeu 51  
Db 246 TGTAGGACATCATCGTGTGCCCTATATTAGGECATTACACAGAACCTC 305

Qy 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspProTrpThrVal 71  
Db 306 AGTTTCCAAAGGGACCAAGGGTGTCTAGGAAATCGGGAGGTGGAAAGGT 365

Qy 72 LeuSerGluValSerGlyArgGluTyrosineProSerValHisValGlyVal -- 90  
Db 366 CCAATCCCTGGCACCCCGAACAGGGCTACTCCAAAGGAACATCGGAACTTC 425

Qy 91 -----SerHisGlyIysValProLeuThrIleValSerGluGlyLeu 107  
Db 426 TCTCTGGAGAACAGGGAGTGGTTTCAGGGCAATCAGCGGGAAAGGACGAGGCCAA 485

Qy 108 LeuLeuLeuProGlyIysValProGlyIysValLeuLeuGluUserGlnThrArgArg 127  
Db 486 CTGCTGGCTCCGGCACACATCTGGCTCTCATGATCGGGATAGCGAGAACCTAA 545

Qy 128 GLYSERTYsSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147  
Db 546 GAAAGTACTCTTGCCCCGAGTACCACTGGAGACTACCGGAAACAT 605

Qy 148 TyrArgIleHisCysteinAspAsnGlyTrpLeuTyrlsSerProArgLeuThrPhePro 167  
Db 606 TACAAGATCCGACCCCTGGAAACGGGGCTCTCATATCCCGAACACCTTCAC 665

Qy 168 SerLeuGlnAlaLeuValAspHistYsSerCysLeuLeu 187  
Db 666 ACTCTGGAGCTGTTGGGACCAACGAGGGCAACACGGGCTCTCGCAGAAACTG 725

Qy 188 LysGluProCysVal 192  
Db 726 TGGTGCCTGCGATG 740

RESULT 11  
US-09-939-853A-75  
Sequence 1983, Application US/0954456  
Patent No. US2002015057A1

GENERAL INFORMATION:  
APPLICANT: Young, Paul  
TITLE OF INVENTION: Sets  
FILE REFERENCE: 689290-76  
CURRENT APPLICATION NUMBER: US/09/954,456  
CURRENT FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/233,617  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/234,052  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234,923  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: US/60/235,134  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: US/60/235,537  
PRIOR FILING DATE: 2000-09-16  
PRIOR APPLICATION NUMBER: US/60/235,638  
PRIOR FILING DATE: 2000-09-26

Alignment Scores:  
Pred. No.: 2.116e-31 Length: 2015  
Score: 360.50 Matches: 77  
Percent Similarity: 58.38% Conservative: 31  
Best Local Similarity: 41.62% Mismatches: 70  
Query Match: 26.64% Indels: 7  
DB: 9 US-09-954-456-1983  
Gaps: 2

US-09-939-853A-75 (1-261) x US-09-954-456-1983 (1-2015)

Qy 12 ProserProSerLeuSerSerValGlnGlyProvalThrMetGluAlaGlu 31  
Db 286 CCGGGCCCTAAAGCCQAACRGCTCTCTAGAGGAAATCAGGGAGGAGGC 336

Qy 32 ArgSerIysValAlaLeuGlySerProAlaGlyGlyProAlaGluLeu 51  
Db 337 TCTGAGGACATATCGGTRGCCCTGTATAGGGCATTACACAGAACCTC 396

Qy 52 SerLeuArgLeuGlyArgGluTyrosineProLeuThrIleValSerGluAspArgIleArgHis 71  
Db 397 AGTTCGAAAGGGAAACAGATGGCTCTAGAGAAATCGGGAGTGTGGCT 456

Qy 72 LeuSerGluValSerGlyArgGluTyrosineProSerValHisValGlyLysVal -- 90  
Db 457 CGATCCCTGGCACCCGGAGGGCTCATCCAAAGCAACTACTGCGCCGCTGAC 516

Qy 91 -----SerHisGlyIysValProGlyIysValLeuSerGlyLeuSerGlyLeu 107  
Db 517 TCTCTGGGACAGGAGGTGGTTTCATAGGCCATAGGGAAAGGACGCCAA 576

Qy 108 LeuLeuLeuProGlyIysValProGlyIysValLeuLeuGluUserGlnThrArgArg 127  
Db 577 CTCCTGGCTCCGGCACACATGTGGCTCATATCCGGATAGGGAGACCCTAA 636

Qy 128 GlySERTYsSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147  
Db 637 GGAGGACTCTTGTGTTGGAGACTACCCGGGGGATACCGTGAACAT 696

Qy 148 TyrArgIleHisCysteinAspAsnGlyTrpLeuTyrlsSerProArgLeuThrPhePro 167  
Db 697 TACAGATCCGACCCCTGACAGCAGGGCTCTGAGAAACCTG 756

Qy 168 SerLeuGlnAlaLeuValAspHistYsSerCysLeuLeu 187  
Db 757 ACCTGCGAGGTGGGACTCTGAGAAACCTG 816

Qy 188 LysGluProCysVal 192  
Db 817 TCCTGGTGCCTGCTGATG 831

RESULT 12  
US-10-007-010-3  
Sequence 3, Application US/10007010  
Publication No. US20030125275A1  
GENERAL INFORMATION:  
APPLICANT: Alexander H. Borchers  
APPLICANT: Kenneth W. Dobie

; TITLE OF INVENTION: ANTISENSE MODULATION OF HCK EXPRESSION  
 ; FILE REFERENCE: RITS-0345  
 ; CURRENT APPLICATION NUMBER: US/10/007,010  
 ; CURRENT FILING DATE: 2001-12-04  
 ; NUMBER OF SEQ ID NOS: 87  
 ; SEQ ID NO: 3  
 ; LENGTH: 2015  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE: CDS  
 ; NAME/KEY: CDS  
 ; LOCATION: (169) . . . (1686)  
 ; US-10-007-010\_3

Alignment Scores:  
 Prod. No.: 2.16e-31  
 Score: 360.50  
 Percent Similarity: 58.62%  
 Best Local Similarity: 41.62%  
 Query Match: 26.64%  
 DB: 14

US-09-939-853a-75 (1-261) × US-10-007-010-3 (1-2015)

Qy 12 ProserProSerLeuSerSerSerValGlnGlyGlnGlyProvalThrMetGluAlaGlu 31  
 Db 286 CCGGCCATATGCCACACAGC-----AACACACAGGAATCAGGGAGCCAGGC 336

Qy 32 ArgSerLysAlaThrAlaValAlaLeuGlySer-ProAlaGlyGlyProAlaGluLeu 51  
 Db 337 TCTGGGACTCATCGTGTGTTGCGCTGTATGAGGCCATTACCAAGAACGCTC 396

Qy 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrrPrrThrVal 71  
 Db 397 AGCTTCCAGAGGGGACAACTCTAGGGATTCAGGATCCGGGAGTGGTGGGAGGCT 456

Qy 72 LeuSerGluValSerGlyArgGluTyrosineProSerValAlaIleProSerValAla 90  
 Db 457 CGAATCCTGGCACCGGAACTGGGGCTACATCCAAAGCAACTATGCGCCGGTGTGAC 516

Qy 91 -----SerHisGlyTrpLeuTyroGluGlyLeuSerArgGluIysAlaGluGluLeu 107  
 Db 517 TCTCTGGGACAGAGGGATGGTGTTCAGGGCATACGGCAAGGAGCGAGGCCCAA 576

Qy 108 LeuLeuLeuProGlyAsnProGlyGlyAlaPhenLeuIleArgGluSerGlnThrArg 127  
 Db 577 CTGGTGGTCCCAGCAATGGGGCTTGATGATGGGATAACGGACCACTAA 636

Qy 128 GlySerTyrosineSerIleSerValArgLeuSerArgProAlaSerTrpAspArgLeuGly 147  
 Db 637 GGAGCTACTCTTGCCGAGACTAGACGACCTCGAGGGAGATACTGTGAAACAT 696

Qy 148 TyrArgIleHisCysLeuAspArgGlyTrpLeuTyrosineSerProAspLeuThrPro 167  
 Db 697 TACAGATCGGACCTGCGAACAGGGGGCTTCTACATATCCCCAAAGCCCTCAGC 756

Qy 168 SerLeuGlnAlaLeuValAspHistidineSerGluLeuAlaAspAspIleCysCysLeu 187  
 Db 757 ACTTGCGGAGAGCTGGTGAACCAATAGAGGGAAACGAGGGCTCTGCGAAACTG 816

Qy 188 LysGluProCysVal 192  
 Db 817 TCGGTGCCCCGTGATC 831

RESULT 13  
 US-10-222-157-140  
 ; Sequence 140, Application US/10252157  
 ; Publication No. US20030190640A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Paris, Mary  
 ; APPLICANT: Pearson, Cecelia I.  
 ; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER  
 ; FILE REFERENCE: PA-0027-1 US

; CURRENT APPLICATION NUMBER: US/10/252,157  
 ; CURRENT FILING DATE: 2002-10-01  
 ; PRIORITY NUMBER: 60/295,048  
 ; PRIORITY FILING DATE: 2001-05-31  
 ; NUMBER OF SEQ ID NOS: 501  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO: 140  
 ; LENGTH: 2341  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No. US20030190640A1 411296.2  
 US-10-252-157-140

Alignment Scores:  
 Prod. No.: 2.65e-31  
 Score: 360.50  
 Percent Similarity: 58.38%  
 Best Local Similarity: 41.92%  
 Query Match: 26.64%  
 DB: 14  
 Gap: 2

US-09-939-853a-75 (1-261) × US-10-252-157-140 (1-2341)

Qy 12 ProserProSerLeuSerSerSerValGlnGlyGlnGlyProvalThrMetGluAlaGlu 31  
 Db 592 CCGGCCATATGCCACACAGC-----AACACACAGGAATCAGGGAGCCAGC 642

Qy 32 ArgSerLysAlaThrAlaValAlaLeuGlySer-ProAlaGlyGlyProAlaGluLeu 51  
 Db 643 TCTGGGACATCAICGTGGTGCCTGATGATACGGGCCATTACACGAGAACCTC 702

Qy 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrrPrrThrVal 71  
 Db 703 AGCTTCCAGAGGGGACCAAGATGTGGTCTAGGATTCGGGAGTGGTGAAGGCT 762

Qy 72 LeuSerGluValSerGlyArgGluTyrosineProSerValAlaIleProSerValAla 90  
 Db 763 CGATCCCTGGCACCGGAACTGGGGCTACATCCAAAGCAACTATGCGCCGGTGTGAC 822

Qy 91 -----SerHisGlyTrpLeuTyroGluGlyLeuSerArgGluIysAlaGluGluLeu 107  
 Db 823 TCTCTGGGAGAGGGAGCTGGTCTAGGGCATACGGGAGGAGCAGGCCAA 882

Qy 108 LeuLeuLeuProGlyAsnProGlyGlyAlaPhenLeuIleArgGluSerGlnThrArg 127  
 Db 883 CTGGTGGCTCCGGCAACATGGGGCTTGATGATGGGATACTGGGAGGATAGGGAGGAACTAA 942

Qy 128 GlySerTyrosineSerIleSerValArgLeuSerArgProAlaSerTrpAspArgLeuGly 147  
 Db 943 GGAACCTACTCTTGCTCCGAGACTAGACGACCTCTGGCAGACTAGACGACCTCTGGCAGACTAA 1002

Qy 148 TyrArgIleHisCysLeuAspArgGlyTrpLeuTyrosineSerProAspLeuThrPro 167  
 Db 1003 TACAGATCGGACCTGCGAACAGGGGGCTTCTACATATCCCCAAAGCCCTCAGC 1062

Qy 168 SerLeuGlnAlaLeuValAspHistidineSerGluLeuAlaAspAspIleCysCysLeu 187  
 Db 1063 ACTTGCGGAGAGCTGGTGAACCAATAGAGGGAAACGAGGGCTCTGCCAGAAACTG 1122

Qy 188 LysGluProCysVal 192  
 Db 1123 TCGGTGCCCCGTGATC 1137

RESULT 14  
 US-10-062-674-2038  
 ; Sequence 140, Application US/10062674  
 ; Publication No. US20040003559A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lorin, Jeanne F.; Kaser, Matthew R.  
 ; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
 ; FILE REFERENCE: PA-0026-1 CIP

CURRENT APPLICATION NUMBER: US/10/062,674  
 CURRENT FILING DATE: 2002-01-30  
 PRIOR APPLICATION NUMBER: US 09/625,102  
 PRIOR FILING DATE: 2000-07-24  
 NUMBER OF SEQ ID NOS: 2217  
 SOFTWARE: PERL Program  
 SEQ ID NO: 2038  
 LENGTH: 2343  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 OTHER INFORMATION: Incyte ID No. US2004000559A1 411296.5

US-09-939-853A-75 (1-261) × US-10-062-674-2038 (1-2343)

Qy    12	ProserProSerIeuSerSerSerValGlnGlyGlnGlyProvalThrMetGluAlaGlu	31
Db    592	CCGGGCCCTATAACCACAAAGC-----AACACACCGAAATCAGGGAGGGGG	642
Qy    32	ArgSerIlysAlaValAlaLeuGlySerPheProAlaGlyGlyProalagluLeu	51
Db    643	TCTAGGACATCATCGTGTGCCTGTATGATTACGGGCCATTACCAAGAACGCTC	702
Qy    52	SerIeuArgLeuGlyGluIeuSerGluAspGlyAspTrpTrpTrpTrpVal	71
Db    703	AGCTTCACAGGGGACCAAGCTGAGCTGCTAGGAATCGGGAGTGTTGGAGGCT	762
Qy    72	LeuSerGluValSerGlyArgGluTyroAsnIleProSerValHisValGlyIysVal---	90
Db    763	CGATCCCTGCCACCGGAAGGGCTACATCCAGAACTATGCGCCGCTGCTGAC	822
Qy    91	-----SerHisGlyTripleutyrGluGlyLeuSerArgGluIysAlaGluIleLeu	107
Db    823	TCTGGGACAACAGGGATGGTTCTAGGGGATCGCCAGAGCGAGGCCAA	882
Qy    108	IeuLeuLeuPrgGlyAsnProGlyGluAlaLeuIleArgGluSerGlnThrArg	127
Db    883	CTGGTGGCTCCGCCAACATGCTGGCTCCATGATCCGGATAGCGAACCTAAA	942
Qy    128	GlySerTySerIeuSerValArgLeuSerArgProIleSerTrpAspar5Ieargin	147
Db    943	GGAAAGCTACTCTGTGCGTAGACTAGACCCCTGGAGGAAATACCGTGAACAT	1002
Qy    148	TyArgIleHisIcysLeuAspHisItySerSerProArgLeuThrPhePro	167
Db    1003	TACAGATCGGCCCTCGAGAACGGGGCTCTACATATCCCCGAAGACCTCAGC	1062
Qy    168	SerIeuGlnAlaLeuIeuAlaAspAspIleCysCysIeuLeu	187
Db    1063	ACTCTGAGGAGTGGGACCACTAACAGAAGGGAACGCGGGCTCTGGCAGAACTG	1122
Qy    188	LysGluProCysVal 192	
Db    1123	TGGTGCCTGCTGCTG 1137	

RESULT 15  
 US-09-917-800A-1611

PATENT NO. US20020119462A1  
 GENERAL INFORMATION:  
 APPLICANT: Mardick, Donna  
 PORTER, Mark  
 JOHNSON, Kory

Qy 114 ProGlyAlaPheLeuIleArgGluSerGlnThrArgGlySerTyrSerLeuSer 133  
 Db 604 ATGCGGGCCTTATGACCCGGACACTGAGCCACAAAGGAGTACTACTCT 663  
 Qy 134 ValArgLeuSerArgProAlaSerTrpAspArgIleArgHistYArgIleHisCysLeu 153  
 Db 664 GTTCAGACTTGACCCCGACGGACCGAGAACCGTGAAGCATTATAATGGACACTG 723  
 Qy 154 AspArgGlyTrpLeuTyrTlleSerProArgLeuThrPhiProSerLeuGlnAlaLeuVal 173  
 Db 724 GACAGTGGAGGGTCTACACTCTCGAGGACCTTCAGGCCCTGCAGGAACATGTC 783  
 Qy 174 AspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysVal-Le 193  
 Db 784 GTCCCACTACAGAAGGGAAAGGGGGCTGTGCAAGAGCTGTCAGTCCCCTGTGTC 843  
 Qy 193 uGin-----ArgAlaGlyProLeuProGlyLysAspIleProLeuProValTh 209  
 Db 844 CGAAACCCAGAACCCATGGAAAAGATGCCCTGGGATTTCTCGAAATCCCTGAG 903  
 Qy 209 rvalGlnArgThr-----ProLeuAsn-TriplysGluLeuAspSerSerLeuPheS 227  
 Db 904 ATGGGAAAGAACTGGGGGGAGTTGGAGAAGTTGGATGGCATGCCACITACAACAG 963  
 Qy 227 ergGluAlaAlaIleThrGlyGluGluSer 235  
 Db 964 CACACAAAGTGCCGTGAGACAT 989

Search completed: February 20, 2004, 12:38:07  
 Job time : 370 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using bw model

Run on: February 19, 2004, 21:21:18 ; Search time 87.562 Seconds  
 (without alignments)  
 9899.970 Million cell updates/sec

Title: US-09-939-853A-140  
 Perfect score: 20  
 Sequence: 1 ctggacaggatctagggttgc 20

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : GenEmbl:\*

	Result No.	Score	Query	Match	Length	DB ID	Description
	1	20	100.0	20	6	AX443199	AX443199 Sequence
	2	20	100.0	1183	6	AX443133	AX443133 Sequence
	3	20	100.0	1183	6	AX443135	AX443135 Sequence
	4	20	100.0	2538	9	BC042041	BC042041 Homo sapi
	5	20	100.0	2567	6	AX452880	AX452880 Sequence
	6	20	100.0	2788	6	AX780857	AX780857 Sequence
	7	20	100.0	66741	9	HS46059	AL021162 Human DNA
	8	20	100.0	145833	2	AC026539	AC026539 Homo sapi
	9	17.4	87.0	80828	9	AC03247	Homo sapi
	10	17.4	87.0	100300	2	AC114970	Homo sapi
	11	17.4	87.0	108742	2	AC028117	Homo sapi
	12	17.4	87.0	110000	2	AC139485_1	Continuation (2 of AC132801) Homo sapi
	13	17.4	87.0	116702	9	AC132801	AC132801 Homo sapi
	14	17.4	87.0	129218	2	AC138851	AC138851 Homo sapi
	15	17.4	87.0	133181	2	AC138930	AC138930 Homo sapi
	16	17.4	87.0	140410	2	AC139483	AC139483 Homo sapi
	17	17.4	87.0	145493	2	AC138839	AC138839 Homo sapi
	18	17.4	87.0	146159	9	AC118459	AC118459 Homo sapi
	19	17.4	87.0	146175	2	AL116133	AL116133 Homo sapi
	20	17.4	87.0	147840	2	AC138939	AC138939 Homo sapi
	21	17.4	87.0	149397	9	AC138827	AC138827 Homo sapi
	22	17.4	87.0	152445	9	AC138410	AC138410 Homo sapi
	23	17.4	87.0	152782	9	AL133255	AL133255 Human DNA
	24	17.4	87.0	154803	9	AC131392	AC131392 Homo sapi
	25	17.4	87.0	155537	2	AC13982	AC13982 Homo sapi
	26	17.4	87.0	156527	2	AC145032	AC145032 Homo sapi
	27	17.4	87.0	158755	2	AC139484	AC139484 Homo sapi
	28	17.4	87.0	159184	2	AC044800	AC044800 Homo sapi
	29	17.4	87.0	160994	2	AC145100	AC145100 Homo sapi
	30	17.4	87.0	161000	2	AC139508	AC139508 Homo sapi
	31	17.4	87.0	161279	2	AC13508	AC13508 Homo sapi
	32	17.4	87.0	161460	9	AC09243	AC09243 Homo sapi
	33	17.4	87.0	161460	9	AL157879	AL157879 Human DNA
	34	17.4	87.0	161671	9	AC140175	AC140175 Homo sapi
	35	17.4	87.0	161893	9	AC137781	AC137781 Homo sapi
	36	17.4	87.0	161957	2	AC138909	AC138909 Homo sapi
	37	17.4	87.0	162703	9	AC132802	AC132802 Homo sapi
	38	17.4	87.0	165211	2	AC145128	AC145128 Homo sapi
	39	17.4	87.0	166615	9	AC138928	AC138928 Homo sapi
	40	17.4	87.0	167663	2	AC145135	AC145135 Homo sapi
	41	17.4	87.0	168547	2	AC138947	AC138947 Homo sapi
	42	17.4	87.0	168814	9	AC010237	AC010237 Homo sapi
	43	17.4	87.0	169082	2	AC138933	AC138933 Homo sapi
	44	17.4	87.0	170336	2	AC011244	AC011244 Homo sapi
	45	17.4	87.0	171363	9	AC145129	AC145129 Homo sapi

## ALIGNMENTS

RESULT	1	LOCUS	AX443199	DEFINITION	Sequence 140 from Parent AX443199	ACCESSION	VERSION	KEYWORDS	synthetic construct
SOURCE									synthetic construct
ORGANISM									artificial sequences
REFERENCE	1	AUTHORS	Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,B.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.	TITLE	Proteins and nucleic acids encoding same				

Pred. No. is the number of results predicted by chance to have a

JOURNAL	Patent: WO 0216599-A 140 28-FEB-2002;	/organism="Homo sapiens"
FEATURES	Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)	/mo_type="unassigned DNA"
Source	1..20	/db_xref="taxon:906"
ORIGIN		
Query Match	Score 20; DB 6; Length 20;	Query Match 100.0%; Score 20; DB 6; Length 1183;
Best Local Similarity	100.0%; Pred. No. 5.2;	Best Local Similarity 100.0%; Pred. No. 4.5;
Matches	20; Conservative 0;	Matches 20; Conservative 0;
Qy	1 CTGGACAGGTTAGGGCTTGT 20	Qy 1 CTGGACAGGTTAGGGCTTGT 20
Db	1 CTGGACAGGTTAGGGCTTGT 20	Db 883 CTGGACAGGTTAGGGCTTGT 902
RESULT 2		RESULT 4
AX443133/c	AX443133	AX443133
LOCUS	1193 bp	1193 bp
DEFINITION	DNA	DNA
SEQUENCE	from Patent WO0216599.	from Patent WO0216599.
VERSION	74	74
KEYWORDS		
ORGANISM	Homo sapiens (human)	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Burgess, C. E., Conley, P. B., Gross, W.M., Hart, M., Kekuda, R., Shimkets, R.A., Spycek, K.A., Szekeres, E.S., Tomlinson, J.E., Topper, J.N. and Yang, R.B.	Burgess, C. E., Conley, P. B., Gross, W.M., Hart, M., Kekuda, R., Shimkets, R.A., Spycek, K.A., Szekeres, E.S., Tomlinson, J.E., Topper, J.N. and Yang, R.B.
TITLE	Proteins and nucleic acids encoding same	Proteins and nucleic acids encoding same
JOURNAL	Patent: WO 0216599-A 74 28-FEB-2002;	Patent: WO 0216599-A 74 28-FEB-2002;
FEATURES	Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)	Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
Source	1..1183	1..1183
ORIGIN		
Query Match	Score 20; DB 6; Length 1183;	Query Match 100.0%; Score 20; DB 6; Length 1183;
Best Local Similarity	100.0%; Pred. No. 4.5;	Best Local Similarity 100.0%; Pred. No. 4.5;
Matches	20; Conservative 0;	Matches 20; Conservative 0;
Qy	1 CTGGACAGGTTAGGGCTTGT 20	Qy 1 CTGGACAGGTTAGGGCTTGT 20
Db	301 CTGGACAGGTTAGGGCTTGT 282	Db 301 CTGGACAGGTTAGGGCTTGT 282
RESULT 3		RESULT 4
AX443135	AX443135	AX443135
LOCUS	1193 bp	1193 bp
DEFINITION	DNA	DNA
SEQUENCE	from Patent WO0216599.	from Patent WO0216599.
VERSION	76	76
KEYWORDS		
ORGANISM	Homo sapiens (human)	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Burgess, C. E., Conley, P. B., Gross, W.M., Hart, M., Kekuda, R., Shimkets, R.A., Spycek, K.A., Szekeres, E.S., Tomlinson, J.E., Topper, J.N. and Yang, R.B.	Burgess, C. E., Conley, P. B., Gross, W.M., Hart, M., Kekuda, R., Shimkets, R.A., Spycek, K.A., Szekeres, E.S., Tomlinson, J.E., Topper, J.N. and Yang, R.B.
TITLE	Proteins and nucleic acids encoding same	Proteins and nucleic acids encoding same
JOURNAL	Patent: WO 0216599-A 76 28-FEB-2002;	Patent: WO 0216599-A 76 28-FEB-2002;
FEATURES	Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)	Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
Source	1..1183	1..1183

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA GI: 28416422.

**FEATURES**  
source  
1. .2538  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:49845 IMAGE:4429896"  
/tissue\_type="Prostate", adenocarcinoma."  
/clone.Lib="NIR MGC\_91"  
/lab\_host="DR10B"  
/note="vector: PCMV-SPORTS"  
1. .2538  
/gene="SLA2"  
/note="SYNonyms: FLJU21992, SLAP-2, C20orf156, MGC49845,  
SLAP2"  
/db\_xref="LocusID:84174"  
/db\_xref="MIM:605577"  
363. .1148

/codon\_start=1  
/product="Src-like-adaptor 2, isoform a"

/protein\_id="AAH42041.1"  
/db\_xref="GI:745943"  
/db\_xref="LocusID:84174"

/translation="MGSLPSREKSLPSPSLSSVCGCGPYTMEEAERSKATAVALGSFP  
AGGPAEELSLRLGPPLIVSEDDWTVLSEGVLSLPSVPAKSHGMLYGEISRS  
KAELLPLLGPNRGAFLIRESDQRGGLPKDIPQFVYQRTPLVTPPLNK  
SPLTFPLQLQDHYELLAIDCCLKEPCVLLQRAGPLPKDIPQFVYQRTPLVTPPLNK  
ELDSLLSEARTEGEESLSEGRSESUSFSYSLNDEAVSLDDA"  
466. .632

misc\_feature  
domains are often indicative of a protein involved in signal transduction related to cytoskeletal organization.  
First described in the Src cytoplasmic tyrosine kinase.  
The structure is a partly opened beta barrel"  
/db\_xref="CDD:pfam0018"  
612. .890

/note="SH3; Region: SH3 (Src homology 3)  
/db\_xref="CDD:pfam0017"  
/db\_xref="SH2; Region: SH2 domain"

**ORIGIN**

Query Match Score 20; DB 9; Length 2538;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 266 CTGGACGGTTAGGGCTTTC 247

**RESULT 5**

AX452880/c  
LOCUS AX452880  
DEFINITION Sequence 1 from Patent WO242457.  
VERSION AX452880  
KEYWORDS  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE Authors Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villar, J.J. and Kanner, S.B.

TITLE Cloning and expression of human slab-2: a novel sh2/sh3 domain-containing human slab homologue having immune cell-specific expression  
JOURNAL Patent: WO 0242457-A 1 30-MAY-2002; Bristol-Myers Squibb Co. (US)

FEATURES source  
1. .2567  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

**FEATURES**  
Location/Qualifiers  
1. .2538

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:49845 IMAGE:4429896"  
/tissue\_type="Prostate", adenocarcinoma."  
/clone.Lib="NIR MGC\_91"  
/lab\_host="DR10B"  
/note="vector: PCMV-SPORTS"

1. .2538  
/gene="SLA2"  
/note="SYNonyms: FLJU21992, SLAP-2, C20orf156, MGC49845,  
SLAP2"  
/db\_xref="LocusID:84174"  
/db\_xref="MIM:605577"

CDS  
/codon\_start=1  
/product="Src-like-adaptor 2, isoform a"

/protein\_id="AAH42041.1"  
/db\_xref="GI:745943"  
/db\_xref="LocusID:84174"

/translation="MGSLPSREKSLPSPSLSSVCGCGPYTMEEAERSKATAVALGSFP  
AGGPAEELSLRLGPPLIVSEDDWTVLSEGVLSLPSVPAKSHGMLYGEISRS  
KAELLPLLGPNRGAFLIRESDQRGGLPKDIPQFVYQRTPLVTPPLNK  
SPLTFPLQLQDHYELLAIDCCLKEPCVLLQRAGPLPKDIPQFVYQRTPLVTPPLNK  
ELDSLLSEARTEGEESLSEGRSESUSFSYSLNDEAVSLDDA"  
363. .1148

/codon\_start=1  
/product="Src-like-adaptor 2, isoform a"  
/protein\_id="AAH42041.1"  
/db\_xref="GI:745943"  
/db\_xref="LocusID:84174"

/translation="MGSLPSREKSLPSPSLSSVCGCGPYTMEEAERSKATAVALGSFP  
AGGPAEELSLRLGPPLIVSEDDWTVLSEGVLSLPSVPAKSHGMLYGEISRS  
KAELLPLLGPNRGAFLIRESDQRGGLPKDIPQFVYQRTPLVTPPLNK  
SPLTFPLQLQDHYELLAIDCCLKEPCVLLQRAGPLPKDIPQFVYQRTPLVTPPLNK  
ELDSLLSEARTEGEESLSEGRSESUSFSYSLNDEAVSLDDA"  
466. .632

misc\_feature  
domains are often indicative of a protein involved in signal transduction related to cytoskeletal organization.

First described in the Src cytoplasmic tyrosine kinase.

The structure is a partly opened beta barrel"  
/db\_xref="CDD:pfam0018"

/note="SH3; Region: SH3 domain"  
/db\_xref="CDD:pfam0017"

**ORIGIN**

Query Match Score 20; DB 9; Length 2538;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 318 CTGGACGGTTAGGGCTTTC 299

**RESULT 6**

AX780857/c  
LOCUS AX780857  
DEFINITION Sequence 3014 from Patent WO03039443.  
VERSION AX780857  
KEYWORDS  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE Authors Haferlach, T., Schoch, C., Kern, W., Kohlmann A., Schnittger, S., Dugas, M., Ellis, R., Brors, B., and Mergenthaler, S.

TITLE JOURNAL Patent: WO 03039443-A 3014 15-MAY-2003; Deutsches Krebsforschungszentrum (DKFZ); Ludwig-Maximilian-Universitaet Muenchen (DB); PD Dr. Dr. Schoch, Claudia (DE); Kern, Wolfgang (DE)

FEATURES source  
1. .2788

/organism="Homo sapiens"  
/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

**RESULT 7**

HS460J8  
LOCUS HS460J8  
DEFINITION Human DNA sequence from clone RP3-460J8 on chromosome 20:11,21-11,23. Contains the 3' end of a gene similar to N-myc downstream regulated (NDRG1) the 5' end of a gene encoding a novel protein tyrosine kinase, ESTs, STSs and GSSs, complete sequence.  
VERSION AL031662  
KEYWORDS  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE Authors Skuce, C.  
TITLE JOURNAL Direct Submission Submitted (23-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
request: clonerques@sanger.ac.uk  
On Aug 7, 2000 this sequence version replaced gi:6425549.

COMMENT During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Bm: EMBL; Sv: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/elegans/wormpep>. This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further details see <http://www.sanger.ac.uk/HGP/Chr20>.

**IMPORTANT:** This sequence is not the entire insert of clone RP3-460J8. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP3-460J8 is at 66741 in this sequence. The true left end of clone RP5-977B1 is at 41767 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP3-460J8 is from the library RPCI-3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>.

## FEATURES

## Source

Location/Qualifiers  
1. . 66741  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="G11\_21\_11\_23"  
/clone="RP3-460J8"  
/clone\_id="RP5-977B1-3"  
complement(50..544;  
/note="match: GSS: Em: B45150"  
complement(240..7128)  
/gene="dJ460J8.1"  
complement(join(<240..339,6995..7128))  
/gene="dJ977B1.1"  
/product="dJ460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"  
/note="match: CDNAs: Em: AK025645  
match: Bm: BG17487;"  
/evidence-not experimental  
complement(join(<240..339,6995..7085))  
/gene="dJ977B1.1"  
/product="Continues in Em: AL050318 as dJ977B1.2"  
/codon\_start=1

/evidence-not experimental  
/product="dJ460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"  
/protein\_id="CA44645.1"  
/db\_xref="GOA:Q9H603"  
/db\_xref="SWISS-PROT:Q9H603"  
/transliteration="MGSLSRRLSPSLSSVQGQGPVITMAERSKATAVALGSFPAGGPAAELSLURGLGEPLITIVS"

## repeat\_region

/note="15 copies 2 mer ag 93% conserved"  
2375..2402

/note="14 copies 2 mer ta 100% conserved"  
2783..3155

/note="match: GSS: Em: A0807191"  
complement(693..7482)

/note="match: GSS: Em: A0556467"  
complement(6985..7492)

/note="match: GSS: Em: A0556478"  
7486..7935

/note="match: STS: Em: A642701S"

repeat\_region

/note="28 copies 2 mer ta 78% conserved"  
34564..34611

## repeat\_region

10615..10640

/note="13 copies 2 mer tt 92% conserved"  
13593..13961

repeat\_region

14248..14297

/note="MERR21B repeat: matches 422..785 of consensus"  
14594..14614

repeat\_region

14746..15078

/note="MERR21B repeat: matches 355..374 of consensus"  
/note="MERR21B repeat: matches 1..355 of consensus"  
complement(17216)

polyA\_site

17519..17686

/note="match: GSS: Em: A2067993"  
complement(17521)

polyA\_site

17522..17523

/note="match: GSS: Em: A2067993"  
complement(17523..54534)

/gene="dJ460J8.3"

complement(join(<17523..19451,20551..20589,222110..222161,  
22247..22282,26087..26134,30796..30847,30986..31051,  
32039..32142,37091..37147,46566..46652,48268..48328,  
50163..50225,53242..53362,54429..54534))

/gene="dJ460J8.3"

/product="dJ460J8.1"

/product=dJ460J8.1 (continued from dJ460J8.1 in  
Em: A0132768)"

/note="match: CDNAs: Em: AB031922 Em: US2073 Em: D67953  
Em: X52845 Em: AF04162 Em: U60593 Em: AF045564 Em: AB333921  
Em: AF159032 Em: AF17402 Em: M9814  
Em: AV0023952 Em: 899263

match: ESTs: Em: AA718726 Em: AA039000 Em: AA445016

Em: AA325826 Em: AA1230982 Em: AA162360 Em: AA162360  
Em: T88705 Em: AA1786673 Em: AA181197 Em: AA764653 Em: AA316771

/evidence-not experimental  
complement(17542..17547)

/gene="dJ460J8.3"

17555..17611

/note="match: STS: Em: G19945"  
join(19202..19451,22110..22282,26087..26127,30796..30847)

/note="match: STS: Em: G23762"  
CDS

complement(join(19270..19451,20551..20589,222110..222161,  
22247..22282,26087..26134,30796..30847,30986..31051,  
32039..32142,37091..37147,46566..46652,48268..48328,  
50163..50225,53242..53362,54429..54534))

/gene="dJ460J8.3"

/note="novel protein (FLWJ13556) similar to N-myc  
downstream regulated (NDRG1)"

match: proteins: Sw: Q62433 Sw: Q92597 Tr: Q922L9 Sw: P97862"

/codon\_start=1

/evidence-not experimental  
Em: A0132768"

/protein\_id="CAB65625.1"

/db\_xref="GI: 6687811"

/db\_xref="GOA:Q9UGV2"

/translation="IEHDISTTHGMVHTTRGLPKGNRPVILTYHDIGLNHKSCFNAAFF  
NFDPQEIRCHPAVCHDAGQQGAPSPTGQYPTMDLAENLPVTLHSLSKII  
GIGVGAGAYLRSRALNHPLVEGSLVNDPCAKGWNIDASLGSUTTVYDILKA  
HHPGEBELQANLDLJOTYPMHADINDQNLQFNSNGRRDIBTEPILGONDNIKS  
KTLKSTLLVGDDNSPAVERVCECSRNPINTLLKADCGGLPQVVIQPGKTEAPK  
YFLQENGTYIYVQLSLHSLSPVSEASMTLARSRTHTSSISLGSGESPPSRVSINQS

DGTQPSCBSDVLDRLHOTMEVSC"

complement(24246..24780)

/gene="dJ460J8.3"

/note="match: GSS: Em: A0592789"

complement(22780..28556)

/gene="dJ460J8.3"

/note="match: GSS: Em: A0308867"

34313..34420

/note="HT1 repeat: matches 1..109 of consensus"  
34505..34560

/note="28 copies 2 mer ta 78% conserved"  
34564..34611

## misc\_feature

10615..10640

/note="13 copies 2 mer tt 92% conserved"  
13593..13961

repeat\_region

14248..14297

/note="MERR21B repeat: matches 422..785 of consensus"  
14594..14614

repeat\_region

14746..15078

/note="MERR21B repeat: matches 355..374 of consensus"  
/note="MERR21B repeat: matches 1..355 of consensus"

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tessyae,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Vieir,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted: (22-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 27, 2000 this sequence version replaced gi:7283243.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: U7115  
 Center clone name: 712\_N\_14  
 Summary Statistics  
 Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 125557 bases at least Q40  
 Consensus quality: 135703 bases at least Q30  
 Consensus quality: 139583 bases at least Q20  
 Insert size: 182000; agarose-fp  
 Insert size: 142233; agarose-fp  
 Quality coverage: 2.6 in Q20 bases; agarose-fp  
 Quality coverage: 3.3 in Q20 bases; sum-of-contigs  
 Quality coverage: 3.3 in Q20 bases; sum-of-contigs  
 \* NOTE: This is a working draft sequence. It currently consists of 37 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
 \* \* \* \* \*  
 1215 1314: contig of 1214 bp in length  
 1315 1314: gap of 100 bp  
 \* \* \* \* \*  
 2673 2673: contig of 1359 bp in length  
 2773 2773: gap of 100 bp  
 \* \* \* \* \*  
 4520: contig of 1747 bp in length  
 4620: gap of 100 bp  
 \* \* \* \* \*  
 5961: contig of 1341 bp in length  
 6061: gap of 100 bp  
 \* \* \* \* \*  
 7719: contig of 1658 bp in length  
 7819: gap of 100 bp  
 7820 9799: contig of 1980 bp in length  
 \* \* \* \* \*  
 9899: gap of 100 bp  
 9900 11434: contig of 1535 bp in length  
 \* \* \* \* \*  
 11534: gap of 100 bp  
 11535 11535: contig of 2848 bp in length  
 \* \* \* \* \*  
 14482: gap of 100 bp  
 14483 14482: contig of 2666 bp in length  
 \* \* \* \* \*  
 17148: contig of 2169 bp in length  
 17248: gap of 100 bp  
 \* \* \* \* \*  
 19339: contig of 1991 bp in length  
 19339 19339: gap of 100 bp  
 \* \* \* \* \*  
 21102: contig of 1763 bp in length  
 21103 21102: gap of 100 bp  
 \* \* \* \* \*  
 23371: contig of 2169 bp in length  
 23372 23471: gap of 100 bp  
 \* \* \* \* \*  
 25782: contig of 2311 bp in length  
 25783 25782: gap of 100 bp  
 \* \* \* \* \*  
 28624: contig of 2942 bp in length  
 28824: gap of 100 bp  
 \* \* \* \* \*  
 28925 28924: contig of 2695 bp in length  
 \* \* \* \* \*  
 31719: gap of 100 bp  
 31720 34679: contig of 2960 bp in length  
 \* \* \* \* \*  
 34680 34779: gap of 100 bp  
 \* \* \* \* \*  
 37527: contig of 2748 bp in length

37528                          \* 37627: gap of 100 bp in length  
                                 \* 40146: contig of 2519 bp in length  
                                 \* 40246: gap of 100 bp in length  
                                 \* 40247: contig of 3497 bp in length  
                                 \* 43743: gap of 100 bp in length  
                                 \* 43744: contig of 3290 bp in length  
                                 \* 47133: gap of 100 bp in length  
                                 \* 47134: contig of 3790 bp in length  
                                 \* 47234: gap of 100 bp in length  
                                 \* 51123: contig of 3812 bp in length  
                                 \* 51124: gap of 100 bp in length  
                                 \* 54936: contig of 4518 bp in length  
                                 \* 59553: gap of 100 bp in length  
                                 \* 59554: contig of 4775 bp in length  
                                 \* 64428: gap of 100 bp in length  
                                 \* 64529: contig of 4683 bp in length  
                                 \* 69211: gap of 100 bp in length  
                                 \* 69212: contig of 3590 bp in length  
                                 \* 69311: gap of 100 bp in length  
                                 \* 72001: contig of 3723 bp in length  
                                 \* 72902: gap of 100 bp in length  
                                 \* 73002: contig of 3723 bp in length  
                                 \* 76725: gap of 100 bp in length  
                                 \* 76825: contig of 4355 bp in length  
                                 \* 81180: gap of 100 bp in length  
                                 \* 81280: contig of 5730 bp in length  
                                 \* 87010: gap of 100 bp in length  
                                 \* 87110: contig of 3746 bp in length  
                                 \* 90856: gap of 100 bp in length  
                                 \* 90556: contig of 5565 bp in length  
                                 \* 96620: gap of 100 bp in length  
                                 \* 102321: contig of 5701 bp in length  
                                 \* 102322: gap of 100 bp in length  
                                 \* 108233: contig of 5772 bp in length  
                                 \* 108294: gap of 100 bp in length  
                                 \* 108394: contig of 8296 bp in length  
                                 \* 116650: gap of 100 bp in length  
                                 \* 116790: contig of 8475 bp in length  
                                 \* 125365: gap of 100 bp in length  
                                 \* 125364: contig of 10990 bp in length  
                                 \* 136355: gap of 100 bp in length  
                                 \* 136354: contig of 9379 bp in length.

FEATURES                          Location/Qualifiers  
                                 1..145833  
                                 /organism="Homo sapiens"  
                                 /mol\_type="genomic DNA"  
                                 /db\_xref="taxon:9606"  
                                 /chromosome="120"  
                                 /map="20"  
                                 /clone="RPCI-11 Human Male BACn  
                                 1..1214  
                                 /note="assembly\_fragment"  
                                 1315..2673  
                                 /note="assembly\_fragment"  
                                 2774..4520  
                                 /note="assembly\_fragment"  
                                 4621..5961  
                                 /note="assembly\_fragment"  
                                 6062..7719  
                                 /note="assembly\_fragment"  
                                 7820..9799  
                                 /note="assembly\_fragment"  
                                 9900..11434  
                                 /note="assembly\_fragment"  
                                 11535..14182  
                                 /note="assembly\_fragment"  
                                 1443..17148  
                                 /note="assembly\_fragment"  
                                 17249..19239  
                                 /note="assembly\_fragment"  
                                 19340..21102  
                                 /note="assembly\_fragment"  
                                 21203..23371  
                                 /note="assembly\_fragment"

Source                                  misc\_feature  
                                 23472..25782  
                                 /note="assembly\_fragment"  
                                 25883..28824  
                                 /note="assembly\_fragment"  
                                 28925..31619  
                                 /note="assembly\_fragment"  
                                 31720..34779  
                                 /note="assembly\_fragment"  
                                 34780..37527  
                                 /note="assembly\_fragment"  
                                 37628..40146  
                                 /note="assembly\_fragment"  
                                 40247..43743  
                                 /note="assembly\_fragment"  
                                 43844..47133  
                                 /note="assembly\_fragment"  
                                 47234..51023  
                                 /note="assembly\_fragment"  
                                 51124..54935  
                                 /note="assembly\_fragment"  
                                 clone\_end:T7  
                                 vector\_side:right"  
                                 55036..59553  
                                 misc\_feature  
                                 Query Match      100.0%      Score 20; DB 2; Length 145833;  
                                 Best Local Similarity 100.0%      Pre<sup>d</sup>. No. 3..9;  
                                 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy                                  1 CTGGACGGTTAGGGTTTG 20  
                                 Db                          100045 CTGGACGGTTAGGGTTG 100026

RESULT 9  
                                 AC093247                          80828 bp      DNA linear PRI 26-FEB-2002  
                                 AC093247                          Homo sapiens chromosome 5 clone RP1-138M1, complete sequence.  
                                 DEFINITION      Homo sapiens chromosome 5 clone RP1-138M1, complete sequence.  
                                 AC093247                          Homo sapiens chromosome 5 clone RP1-138M1, complete sequence.  
                                 AC093247.3                          GI:18921295  
                                 HTG.                          Homo sapiens (human)  
                                 AC093247                          Homo sapiens  
                                 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
                                 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
                                 REFERENCE      1 (bases 1 to 80828)  
                                 AUTHORS                  DOE Joint Genome Institute and Stanford Human Genome Center.  
                                 TITLE                          Direct Submission  
                                 JOURNAL                          Unpublished  
                                 COMMENT                          3 (bases 1 to 80828)  
                                 REFERENCE      2 (bases 1 to 80828)  
                                 AUTHORS                  DOE Joint Genome Institute and Stanford Human Genome Center.  
                                 TITLE                          Direct Submission  
                                 JOURNAL                          Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
                                 REFERENCE      3 (bases 1 to 80828)  
                                 AUTHORS                  DOE Joint Genome Institute and Stanford Human Genome Center.  
                                 TITLE                          Direct Submission  
                                 JOURNAL                          Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
                                 COMMENT                          On Feb. 26, 2002 this sequence replaced gi:15290461.  
                                 REFERENCE      Draft Sequence Produced by DOE Joint Genome Institute  
                                 www.Jgi.doe.gov  
                                 Finishing Completed at Stanford Human Genome Center  
                                 www.SBGC.stanford.edu  
                                 Quality: Phred Quality >=40 99.5% of Sequence;  
                                 Estimated Total Number of Errors is 0.4.  
                                 NOTE: This insert is not the entire sequence of the clone (entire  
                                 sequence is 147.2kb). It is clipped at the overlap with AC010237.  
                                 The number of bases overlapped is 38259.  
                                 Location/Qualifiers  
                                 Location/Qualifiers  
                                 1..80828  
                                 /organism="Homo sapiens"  
                                 /mol\_type="genomic DNA"  
                                 /db\_xref="taxon:9606"

/chromosome="5"  
 /clone="RP11-138M1"

**ORIGIN**

Query Match Score 17.4; DB 9; Length 80828;  
 Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGCTT 19  
 Db 71309 CTGGACAGGTAGGCTT 71327

**RESULT 10**  
**AC14970** AC114970 100300 bp DNA linear HTG 14-MAR-2002  
**DEFINITION** Homo sapiens chromosome 5 clone RP11-388B4, WORKING DRAFT SEQUENCE,  
 7 unordered pieces.

**AC14970** AC114970.1 GI:19424431  
**KEYWORDS** HTGS\_PASEL; HTGS\_DRAFT; HTGS\_ACTIVEFIN.

**ORGANISM** Homo sapiens (human)

**Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Etheria; Primates; Catarrhini; Hominidae; Homo.**

**REFERENCE** 1 (bases 1 to 100300);  
**AUTHORS** DOE Joint Genome Institute.  
**JOURNAL** Unpublished sequencing of Human Chromosome 5  
 2 (bases 1 to 100300);  
**REFERENCE** 2 (bases 1 to 100300);  
**AUTHORS** DOE Joint Genome Institute.  
**JOURNAL** Direct Submission

**COMMENT** Submitted (14-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 -----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: <http://www.jgi.doe.gov>

-----Project Information  
 Center Project Name: 553385  
 Center clone name: RP11-138B4

-----Summary Statistics

Consensus quality: 95359 bases at least Q40  
 Consensus quality: 97011 bases at least Q30  
 Consensus quality: 97025 bases at least Q20  
 Estimated insert size: 175000; agarose-fp estimation  
 Estimated insert size: 99700; sum-of-contigs estimation  
 Quality coverage: 9.62 in Q20 bases; agarose-fp estimation  
 Quality coverage: 16.88 in Q20 bases; sum-of-contigs estimation

\* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known, and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 1595: contig of 1595 bp in length  
 \* 1695: gap of unknown length  
 \* 1696: contig of 1499 bp in length  
 \* 3195: 3294: gap of unknown length  
 \* 3295: 7698: contig of 4404 bp in length  
 \* 7699: 7798: gap of unknown length  
 \* 18801: 18901: contig of 11003 bp in length  
 \* 18802: 18902: 32519: gap of unknown length  
 \* 32520: 32620: 37555: contig of 4976 bp in length  
 \* 37596: 37655: gap of unknown length  
 \* 37696: 100300: contig of 62605 bp in length

**FEATURES** Location/Qualifiers

source 1. .100300  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="RP11-388B4"  
 /clone\_lib="RPCI human BAC library 11"

ORIGIN

Query Match Score 17.4; DB 2; Length 100300;  
 Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGCTT 19  
 Db 64854 CTGGACAGGTAGGCTT 64872

**RESULT 11**  
**AC023817** AC023817 Homo sapiens chromosome 5 clone CTB-2565N12, WORKING DRAFT  
**DEFINITION** Homo sapiens chromosome 5 ordered pieces.

**AC023817** AC023817.5 GI:19224766  
**VERSION** HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_ACTIVEFIN.

**ORGANISM** Homo sapiens (human)

**Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Etheria; Primates; Catarrhini; Hominidae; Homo.**

**REFERENCE** 1 (bases 1 to 10742)  
**AUTHORS** DOE Joint Genome Institute.  
**JOURNAL** Sequencing of Human Chromosome 5  
 Unpublished

**REFERENCE** 2 (bases 1 to 106742)  
**AUTHORS** DOE Joint Genome Institute.  
**JOURNAL** Direct Submission

**COMMENT** On Mar 7, 2002 this sequence version replaced gi:18139273.

-----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: <http://www.jgi.doe.gov>

-----Project Information  
 Center Project Name: 815265  
 Center clone name: CTB-El\_2565N12

-----Summary Statistics

Consensus quality: 107872 bases at least Q40  
 Consensus quality: 108273 bases at least Q30  
 Consensus quality: 108423 bases at least Q20  
 Estimated insert size: 137000; agarose-fp estimation  
 Estimated insert size: 108512; sum-of-contigs estimation  
 Quality coverage: 8.13 in Q20 bases; agarose-fp estimation  
 Quality coverage: 10.27 in Q20 bases; sum-of-contigs estimation.

\* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

\* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

\* the accession number will be preserved.

\* 31413: contig of 31413 bp in length  
 \* 31414: 31513: gap of unknown length

\* 31514 101416: contig of 69903 bp in length  
 \* 101516: gap of unknown length.  
**FEATURES**  
**source** Location/Qualifiers  
 1..108742 Contig of 7226 bp in length.  
**organism** "Homo sapiens"  
**/mol\_type**="genomic DNA"  
**/db\_xref**="taxon:9606"  
**/chromosome**=15"  
**/clone**="CTD-565N12"  
**/clone\_lib**="CalTech human BAC library D"

**ORIGIN**

Query Match	Score 17.4;	DB 2;	Length 108742;
Best Local Similarity	94.7%;	Pred. No. 1.2e+02;	
Matches	18;	Conservative	0;
Qy	1 CTGGACAGCTTAGGGCTTT 19		
Db	70123 CTGGACAGCTTAGGGCTTT 70141		

**RESULT 12**  
**AC139485\_1**  
**WPCOMMENT**  
 Sequence split into 4 fragments LOCUS AC139485 Accession AC139485  

Fragment Name	Begin	End
AC139485_0	1	110000
AC139485_1	100001	210000
AC139485_2	310000	320000
AC139485_3	320000	368755

 Continuation (72 of 4) of AC139485 from base 100001 (AC139485 Homo sapiens chromosome 5 c

Query Match	Score 17.4;	DB 2;	Length 110000;
Best Local Similarity	94.7%;	Pred. No. 1.2e+02;	
Matches	18;	Conservative	0;
Qy	1 CTGGACAGCTTAGGGCTTT 19		
Db	29275 CTGGACAGCTTAGGGCTTT 29293		

**RESULT 13**  
**AC132801\_1**  
**LOCUS** Homo sapiens chromosome 5 clone CTC-202015, complete sequence.  
**DEFINITION** Homo sapiens chromosome 5 clone CTC-202015, complete sequence.  
**ACCESSION** AC132801  
**VERSION** AC132801.2  
**KEYWORDS** HTG.  
**ORGANISM** Homo sapiens (human)  
**Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.**  
**REFERENCE** 1 (bases 1 to 116702)  
**AUTHORS** DOE Joint Genome Institute and Stanford Human Genome Center.  
**JOURNAL** Direct Submission  
**TYPE** Unpublished  
**REFERENCE** 2 (bases 1 to 116702)  
**AUTHORS** DOE Joint Genome Institute.  
**JOURNAL** Direct Submission  
**REFERENCE** 3 (bases 1 to 116702)  
**AUTHORS** DOE Joint Genome Institute and Stanford Human Genome Center.  
**JOURNAL** Direct Submission  
**REFERENCE** 4 (bases 1 to 116702)  
**AUTHORS** DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
**JOURNAL** Direct Submission  
**COMMENT** On Nov 13, 2002 this sequence replaced gi:22711579.  
**REFERENCE** Draft Sequence Produced by DOE Joint Genome Institute  
[www.jgi.doe.gov](http://www.jgi.doe.gov)  
**FINISHING** Completed at Stanford Human Genome Center  
**Quality**: Phrap Quality >=40 100% of Sequence;

Estimated Total Number of Errors is 0.  
**FEATURES**  
**source** Location/Qualifiers  
 1..116702  
**organism** "Homo sapiens"  
**/mol\_type**="genomic DNA"  
**/db\_xref**="taxon:9606"  
**/chromosome**=15"  
**/clone**="CTC-202015"

**ORIGIN**

Query Match	Score 17.4;	DB 9;	Length 116702;
Best Local Similarity	94.7%;	Pred. No. 1.2e+02;	
Matches	18;	Conservative	0;
Qy	1 CTGGACAGCTTAGGGCTTT 19		
Db	60842 CTGGACAGCTTAGGGCTTT 60824		

**RESULT 14**  
**AC138851/c**  
**LOCUS** AC138851  
**DEFINITION** Homo sapiens chromosome 5 clone RP11-1310P17, WORKING DRAFT  
**SEQUENCE**, 13 unordered pieces.

**ACCESSION** AC138851  
**VERSION** AC138851.1 GI:27805263  
**KEYWORDS** HTG; HTGS PHASE; HTGS\_DRAFT; HTGS\_ACTIVEFIN.

**ORGANISM** Homo sapiens  
**SOURCE** Homo sapiens (human)  
**Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.**

**REFERENCE** 1 (bases 1 to 12218)  
**AUTHORS** DOE Joint Genome Institute.  
**TITLE** Sequencing of Human Chromosome 5  
**JOURNAL** Unpublished  
**COMMENT** Center: Joint Genome Institute  
**Center Code:** JGI  
**Web site:** <http://www.jgi.doe.gov>

**Project Information**  
**Center Project Name:** RP11-1310P17  
**Center clone name:** RP11-1310P17

**Summary Statistics**

Consensus quality: 124109 bases at least Q40	-----
Consensus quality: 125106 bases at least Q30	-----
Consensus quality: 125593 bases at least Q20	-----
Estimated insert size: 175000; agarose-fp estimation	-----
Estimated insert size: 128018; sum-of-contigs estimation	-----
Quality coverage: 8.8 in Q20 bases; sum-of-contigs estimation	-----
Quality coverage: 11.03 in Q20 bases; sum-of-contigs estimation	-----
* NOTE: This is a working draft, sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.	-----
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	-----
* 1551: contig of 1550 bp in length	-----
* 1551: gap of unknown length	-----
* 1651: contig of 6079 bp in length	-----
* 7720: gap of unknown length	-----
* 7730: contig of 1968 bp in length	-----
* 7830: 9797: contig of 9899 bp of unknown length	-----
* 9798: 9898: 12555: contig of 2662 bp in length	-----
* 12650: 12659: gap of unknown length	-----

Quality Coverage: 18.4 in Q20 bases; sum-of-contigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently  
   \* consists of 5 contigs. The true order of the pieces  
   \* is not known and their order in this sequence record is  
   \* arbitrary. Gaps between the contigs are represented as  
   \* runs of N, but the exact sizes of the gaps are unknown.  
   \* This record will be updated with the finished sequence  
   \* as soon as it is available and the accession number will  
   \* be preserved.

<p>FEATURES</p> <p>Source</p> <p>ORGIN</p> <p>Query Match Score 17.4; DB 2; Length 133181;    Best Local Similarity 94.7%; Pred. No. 1.2e+02;    Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p> <p>QY 1 CTGGACAGCTTAGGGCTTT 19</p> <p>Db 124866 CTGGACAGCTTAGGGCTTT 124848</p> <p>RESULT 15</p> <p>AC138930 AC138930 133181 bp linear HTG 21-JAN-2003</p> <p>DEFINITION Homo sapiens chromosome 5 clone RP1-6_78014, WORKING DRAFT</p> <p>SEQUENCE, 5 unordered pieces.</p> <p>AC138930 AC138930.1 GI:27805342</p> <p>VERSION HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.</p> <p>KEYWORDS</p> <p>SOURCE</p> <p>ORGANISM</p> <p>BioProject</p> <p>BioProject ID: 133181</p> <p>REFERENCE</p> <p>AUTHORS</p> <p>JOURNAL</p> <p>REFERENCE 2 (bases 1 to 133181)</p> <p>AUTHORS</p> <p>TITLE</p> <p>JOURNAL</p> <p>COMMENT</p> <p>DOE Joint Genome Institute</p> <p>Unpublished sequencing of Human Chromosome 5</p> <p>Center: Joint Genome Institute</p> <p>Center Code: JGI</p> <p>Web site: <a href="http://www.jgi.doe.gov">http://www.jgi.doe.gov</a></p> <p>Project Information</p> <p>Center Project Name: 1555206</p> <p>Center clone name: RP1-11_678014</p> <p>-----</p> <p>Summary Statistics</p> <p>Consensus Quality: 131745 bases at least Q40</p> <p>Consensus Quality: 111928 bases at least Q30</p> <p>Consensus Quality: 131052 bases at least Q20</p> <p>Estimated insert size: 17500; agarose-fp estimation</p> <p>Estimated insert size: 132781; sum-of-contigs estimation</p> <p>Quality Coverage: 13.96 in Q20 bases; agarose-fp estimation</p>
---

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:15:48 ; Search time 18.5392 Seconds  
 4585.415 Million cell updates/sec

Title: US-09-939-853A-140  
 Perfect score: 20  
 Sequence: 1 ctggacaggtaggcttg 20

Scoring table: IDENTITY NUC  
 Gap0 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : N\_Geneseq\_29Jan04;\*

- 1: Geneseqn180s;\*
- 2: Geneseqn1990s;\*
- 3: Geneseqn2000s;\*
- 4: Geneseqn2001aa;\*
- 5: Geneseqn2001bb;\*
- 6: Geneseqn2002s;\*
- 7: Geneseqn2003aa;\*
- 8: Geneseqn2003bb;\*
- 9: Geneseqn2003cc;\*
- 10: Geneseqn2004a;\*

Pred. No. is the number of results Predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	100.0	20	6 ABK61505	Abk61505 Human ORF
C 2	20	100.0	444	6 ABQ98669	Abq98669 Human ORF
C 3	20	100.0	445	5 AA\$74747	Aas74747 DNA encod
C 4	20	100.0	763	6 ABQ98670	Abq98670 Human ORF
C 5	20	100.0	1183	6 ABK61465	Abk61465 Human ORF
C 6	20	100.0	2567	6 ADD13980	Add13980 Human Src
C 7	16.8	84.0	538	9 ADD22694	Add22694 Filamento
C 8	16.8	84.0	589	7 ABZ56309	Abz56309 Aspergil
C 9	16.8	84.0	809	3 AA\$14531	Aas14531 Aspergil
C 10	20	100.0	2305	7 ADA53227	Ada53227 Human Cod
C 11	16.8	84.0	4325	8 ADA53960	Ada53960 Human DAD
C 12	16.8	84.0	4325	9 ADBY2698	Adb72698 Human DAD
C 13	16.8	84.0	4325	9 ADCB84440	Adcb84440 Human DAD
C 14	16.4	82.0	2446	9 ADE57896	Ade57896 Human Gen
C 15	16.4	82.0	2446	9 ADE57892	Ade57892 Human Gen
C 16	16.4	82.0	2446	9 ADE57900	Ade57900 Human Gen
C 17	16.4	82.0	2446	9 ADE57904	Ade57904 Human Gen
C 18	16.4	82.0	2446	9 ADE83385	Ade83385 Human Gen
C 19	15.8	79.0	403	8 ACH8692	Ach8692 Human leu
C 20	15.8	79.0	457	8 ACHA0016	Acha0016 Human foe
C 21	15.8	79.0	712	5 ACDB5832	Acdb5832 Human col
C 22	15.8	79.0	810	5 ABA16524	Aba16524 Human met
C 23	15.8	79.0	1152	2 AXI4434	Axi4434 H. pylori

## ALIGNMENTS

c	24	15.8	79.0	1255	9 ADE56623
c	25	15.8	79.0	1255	9 ADE56627
c	26	15.8	79.0	1255	9 ADE56619
c	27	15.8	79.0	1255	9 ADE56619
c	28	15.8	79.0	1274	2 AA\$42126
c	29	15.8	79.0	3153	2 AAV06254
c	30	15.8	79.0	3362	2 AAV06254
c	31	15.8	79.0	3747	2 AAV06250
c	32	15.8	79.0	4049	4 AAR02878
c	33	15.8	79.0	4170	4 AAD09546
c	34	15.8	79.0	4170	10 ADE77056
c	35	15.8	79.0	24053	4 AAD1308
c	36	15.8	79.0	32768	2 AAX20515
c	37	15.4	77.0	673	6 ABK63392
c	38	15.4	77.0	673	9 ADB57753
c	39	15.4	77.0	2067	4 AAF58597
c	40	15.4	77.0	2067	6 ABZ35294
c	41	15.4	77.0	3685	6 AAS98077
c	42	15.4	77.0	3685	7 ABZ42564
c	43	15.4	77.0	4112	5 Aah81792
c	44	15.4	77.0	4137	3 AAC76337
c	45	15.4	77.0	6029	4 ARK82712

XX	XX	XX	XX	XX	XX
ID	ABK61505	standard	DNA;	20 BP.	
AC	ABK61505;				
DT	18-JUN-2002	(first entry)			
DB	Human NOV13 RT-PCR primer #1.				
HW	Human: ss; NOVX; gene therapy; cardiomopathy; atherosclerosis; cell signal processing disorder; metabolic pathway modulation disorder; diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; primer; uterus cancer; immune response; graft-versus-host disease; exon linking; acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease; hypertension; congenital heart defects; multiple sclerosis; inflammation; Albright hereditary osteodystrophy; reverse transcriptase PCR.				
OS	Homo sapiens.				
XX	XX	XX	XX	XX	XX
PN	WO200216599-A2.				
PD	28-FEB-2002.				
PP	27-AUG-2001; 2001W0-US026510.				
PR	25-AUG-2000; 2000US-0228191P.				
PR	08-FEB-2001; 2001US-0267300P.				
PR	20-FEB-2001; 2001US-0265961P.				
PR	20-MAR-2001; 2001US-0277337P.				
XX	(CURA-) CURAGEN CORP.				
XX	(COR-) COR THERAPEUTICS INC.				
XX	New polypeptides for treating or preventing a disorder associated with them, in humans, e.g. cardiomopathy, atherosclerosis or cancers.				
PI	Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA, Spytak KA, Szekeres ES, Tomlinson JB, Topper JN, Yang R;				
XX	Example 2: Page 234; 263pp; English.				
PS	The invention relates to an isolated polypeptide (NOVX) a mature form of NOVX, a NOVX variant (differing by no more than 15%), the nucleotide CC				

encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14, CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it, CC and antibody against it, are useful for treating or preventing (e.g. by gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy, CC atherosclerosis, a disorder related to cell signal processing and metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide CC and nucleic acids are also useful for determining the presence of predisposition to the disease. The NOVX nucleic acid and polypeptide are especially useful in therapeutic or prophylactic applications for disorders associated with aberrant NOVX expression or activity, e.g. cancer, adenocarcinoma, lymphoma, prostate cancer or uterus, e.g. cancer, immune response, graft-versus-host disease, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension, congenital heart defects, multiple sclerosis, inflammation or Albright hereditary osteodystrophy and many other diseases listed in the specification. The DNA encoding the protein is useful in gene therapy for chromosome mapping, tissue typing, diagnostic or prognostic assays, or for developing a powerful assay system for functional analysis of various human disorders, as well as in diagnostic applications. The present sequence is a reverse transcriptase (RT)-PCR primer used to measure tissue specific expression of mRNA encoding a NOVX protein.

Sequence 20 BP; 3 A; 3 C; 8 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7; Matches 0; Indels 0; Gaps 0;

Db 1 CTGGACAGGTAGGGCTTTCG 20  
Qy 1 CTGGACAGGTAGGGCTTTCG 20

RESULT 2  
ABQ98669/C  
ID ABQ98669 standard; DNA; 444 BP.

XX AC ABQ98669;

XX DT 04-NOV-2002 (first entry)

XX DE Human ORF476 coding sequence.

XX CYTOSTATIC; CARDIANT; ANTI-ALLERGIC; IMMUNOSUPPRESSIVE; VULNERARY; ANTIINFLAMMATORY; GENE THERAPY; HUMAN; ORF8X; ATHEROGENIC; PLATELET; HUMAN UMBILICAL VEIN ENDOTHELIAL CELL; HUVEC; ATHEROCEROTIC PLAQUE; CANCER; CARDIOVASCULAR DISEASE; ALLERGY; AUTOIMMUNE DISEASE; WOUND HEALING; BLOOD COAGULATION DISORDER; INFLAMMATORY DISORDER; DS.

XX OS Homo sapiens.

XX PN US2002082206-A1.

XX PD 27-JUN-2002.

XX PP 30-MAY-2001; 2001US-00867550.

XX PR 30-MAY-2006; 2006US-0208427P.

XX XX (LEAC') LEACH N. D.

PA PA (MEHR') MEHRABAN F.

PA (CONL') CONLEY P. B.

PA (TOPP') TOPPER J. N.

PA (LAWD') LAW D.

XX PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

XX DR WPI; 2002-626554/67.

DR P-PSDB; AB964106.

XX New polypeptide designated ORFX are present in human atherogenic cells and are useful to prevent and treat ORFX-associated disorders including PT and PT

CC cancer, allergy, wound healing or autoimmune, cardiovascular or inflammatory disease.

CC Claim 2; SEQ ID NO 951; 78pp; English.  
PR XX  
PS XX

CC The present invention relates to novel human ORFX polypeptides and their CC coding sequences (AB6361-AB64681 and ABQ981-94 ABQ9267). The sequences CC were discovered in human atherogenic cells, in particular in platelets CC and human umbilical vein endothelial cells (HUVEC) and are expressed in CC many other tissues as well. Atherogenic cells are cells which have the CC potential to develop atherosclerotic plaques. The ORFX polypeptides and CC nucleic acids are useful for treating or preventing a pathological CC condition associated with an ORFX-associated disorder, e.g. cancer, CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood CC coagulation disorders or inflammatory disorders. Note: The sequence data CC for this patent did not form part of the printed specification, but was CC obtained in electronic format directly from the USPTO web site at CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
XX Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 444;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGCTTTCG 20  
Db 60 CTGGACAGGTAGGGCTTTCG 41

RESULT 3  
AAS74747;  
ID AAS74747 standard; cDNA; 445 BP.  
XX AC AAS74747;  
XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #10551.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX DT 11-OCT-2001.  
XX DE 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSEQ ) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX PR WPI; 2001-639362/73.  
DR P-PSDB; ABG1056.  
XX

CC New isolated polynucleotide and encoded polypeptides, useful in PT diagnostics, forensics, gene mapping, identification of mutations PR responsible for genetic disorders or other traits and to assess PT biodiversity.

CC XX Claim 1; SEQ ID NO 10551; 103pp; English.  
PS XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) CC sequences. (I) is useful as hybridisation probes, polymerase chain CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping, CC and in recombinant production of (II). The polynucleotides are also used CC in diagnostics as expressed sequence tags for identifying expressed

genes. (II) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing protein (II) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS641972AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct-Sequences](ftp://wipo.int/pub/published_pct Sequences)

Sequence 445 BP; 89 A; 112 C; 143 G; 101 T; 0 U; 0 Other;

Query 1 CTGGACAGGTAGGGCTTG 20  
Best Local Similarity 100.0%; Score 20; DB 5; Length 445;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 168 CTGGACAGGTAGGGCTTG 187

RESULT 4  
ABQ98670/c  
ID ABQ98670 standard; DNA; 763 BP.  
XX 04-NOV-2002 (first entry)  
AC ABQ98670;  
XX Human ORF477 coding sequence.  
XX Cyrostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary; human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque; cancer; cardiovascular disease; allergy; autoimmune disease; wound healing; blood coagulation disorder; inflammatory disorder; ds.  
XX Homo sapiens.  
OS US2002082206-A1.  
XX 27-JUN-2002.  
XX 30-MAY-2001; 2001US-00867550.  
PR 30-MAY-2000; 2000US-0208427P.  
PA (LEAAC) LEACH M. D.  
PA (MEHR/) MEHRABAN P.  
PA (CONL/) CONLEY P. B.  
PA (TOPP/) TOPPER J. N.  
PA (LAWD/) LAW D.  
PI Leach MD, Mehraban P, Conley PB, Topper JN, Law D;  
XX WPI; 2002-626554/67.  
DR P-PDB; ABP64107.  
XX New polypeptide designated ORFX are present in human atherogenic cells and are useful to prevent and treat ORFX-associated disorders including cancer, allergy, wound healing or autoimmune, cardiovascular or inflammatory disease.

PT The present invention relates to novel human ORFX polypeptides and their  
PT SEQ ID NO 953; 78PP; English.  
PS The invention relates to novel human ORFX polypeptides and their  
CC coding sequences (ABP61631-ABP64681 and ABQ98194-ABQ99267). The sequences  
CC were discovered in human atherogenic cells, in particular in platelets  
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
CC many other tissues as well. Atherogenic cells which have the  
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
CC nucleic acids are useful for treating or preventing a pathological  
CC condition associated with an ORFX-associated disorder, e.g. cancer,  
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
CC coagulation disorders or inflammatory disorders. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from the USPTO web site at  
CC seqdt.uspto.gov/sequence.html?docID=20020082206  
SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;  
SQ Query Match 100.0%; Score 20; DB 6; Length 763;  
SQ Best Local Similarity 100.0%; Pred. No. 2-4;  
SQ Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
SQ Db 189 CTGGACAGGTAGGGCTTG 170

RESULT 5  
ABK61465/c  
ID ABK61465 standard; cDNA; 1183 BP.  
XX AC ABK61465;  
XX DT 18-JUN-2002 (first entry)  
XX Human cDNA encoding protein NOV13.  
XX Human; gene; ss: NOVX; gene therapy; cardiomyopathy; atherosclerosis;  
KW cell signal processing disorder; metabolic pathway modulation disorder;  
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;  
KW uterus; cancer; immune response; graft-versus-host disease;  
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;  
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;  
KW Albright hereditary osteodystrophy.  
XX OS Homo sapiens.  
PN WO200216599-A2.  
XX PD 28-FEB-2002.  
XX PP 27-AUG-2001; 2001WO-US026510.  
XX PR 25-AUG-2000; 2000US-0228191P.  
PR 08-FEB-2001; 2001US-0267300P.  
PR 20-FEB-2001; 2001US-0263961P.  
PR 20-MAR-2001; 2001US-0277337P.  
XX PA (CURA-) CURAGEN CORP.  
PA (CORT-) COR THERAPEUTICS INC.  
EI Burgess CE, Conley PB, Grossie WM, Hart M, Kekuda R, Shimkets RA;  
EI Spytek KA, Sekeris ES, Tomlinson JE, Topper JN, Yang R;  
DR WPI; 2002-280937/32.  
DR P-PSDB; AAU91308.  
XX PT New polypeptides for treating or preventing a disorder associated with  
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.  
XX FS Claim 1; Page 98; 263PP; English.  
XX CC The invention relates to an isolated polypeptide (NOVX) a mature form of  
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide  
CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it

CC and antibody against it, are useful for treating or preventing (e.g., by  
 CC gene therapy) a NOX-associated disorder in human, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 predisposition to the diseases. The NOX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation or Albright  
 CC hereditary osteodystrophy and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence encodes a NOX protein.

XX Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGCTTG 20  
 Db 301 CTGGACAGGTAGGGCTTG 282

RESULT 6  
 ADD43980/C  
 ID ADD43980 standard; CDNA; 2567 BP.

AC ADD43980;

XX DT 13-DEC-2002 (first entry)

DE Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.

XX Human: SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation; disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective; gene; ss.  
 XX OS Homo sapiens.

XX Key Location/Qualifiers  
 FH 415..1200  
 FT /\*tag= a  
 FT /product= "Human SLAP-2"  
 XX WO200242457-A1.  
 XX PD 30-MAY-2002.  
 XX PF 20-NOV-2001; 2001WO-US043367.  
 XX PR 22-NOV-2000; 2000US-0252545P.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX PA Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB,  
 XX DR WPI: 2002-463632/49.  
 XX DR -PSDB; AAE26357.

XX Novel substantially purified human SH2/SH3-domain-containing adapter  
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic  
 PT protein.

PT intervention in immunological and inflammatory disorders and cancer.  
 XX Claim 2; Fig 1; 85PP; English.

XX The invention relates to a substantially purified human SH2/SH3-domain-containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-2). The invention is useful for treating an immune disorder involving hyperactivity of B- or T-lymphocytes in a mammal. The invention is useful for screening for antagonists or inhibitors of the interaction of hSLAP-2 with cellular signalling compounds, for diagnosing, treating or preventing diseases or disorders associated with aberrant or uncontrolled cellular signal transduction, for determining those cellular signalling molecules which associate with hSLAP-2 and which provide critical signals for cell activation, and as effectors in methods to affect T-cell activation. The invention is useful in screening assays to identify and detect candidate biactive agents that modulate hSLAP-2 biactivity, for potential use to treat autoimmune diseases which may be caused by hyperactivated T cells, as well as to treat diseases which may be caused by hyperactivated B cells, in addition to other immune system related conditions, diseases, or disorders, T-cell and B-cell neoplasms, inflammation disorders, diseases and conditions, rheumatoid arthritis, osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's and ulcerative colitis), allergies, particularly those involving hyperactivity of B-cells and T-cells, or other immune cells, such as mast cells or eosinophils, autoimmune diseases such as systemic lupus erythematosus and multiple sclerosis, pulmonary diseases including asthma, acute respiratory distress syndrome, and chronic obstructive pulmonary disorder, tissue/organ rejection, and cancer. The invention is useful in gene therapy. The present sequence is human SLAP-2 cDNA.

SQ Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 2567;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTGGACAGGTAGGGCTTG 20  
 Db 318 CTGGACAGGTAGGGCTTG 299

RESULT 7  
 ADD22694/C  
 ID ADD22694 standard; DNA; 538 BP.  
 XX ADD22694;  
 AC AC  
 XX DT 15-JAN-2004 (first entry)  
 XX DB 318 CTGGACAGGTAGGGCTTG 299  
 XX Filamentous fungi Aspergillus genus DNA sequence, SEQ ID No 145.  
 KW promoter activity; filamentous fungi Aspergillus genus;  
 KW transcription activation; Glucose; gene transcription; ds  
 XX OS Aspergillus oryzae.  
 XX PN JP2003144171-A.  
 XX PD 20-MAY-2003.  
 XX PF 16-NOV-2001; 2001JP-00351368.  
 XX PR 16-NOV-2001; 2001JP-00351368.  
 XX PA (HGET ) HIGETA SHOYU KK.  
 PA (DOKU-) DOKURITSU GIOSEI HOJIN SANGYO GIJUTSU SO.  
 XX DR WPI; 2003-818168/77.  
 XX PT Novel promoter DNA derived from Aspergillus genus, useful for protein  
 PT production.  
 XX PS Claim 1; SEQ ID NO 145; 83PP; Japanese.

XX The invention relates to a novel DNA sequence comprising any one of 74 sequences with promoter activity, having a fully defined sequence shown in the specification or having a sequence which hybridised under stringent conditions to one of the 74 sequences. The invention further comprises one of the 74 DNA sequences or their DNA fragments, derived from filamentous fungi Aspergillus genus, being useful for manufacturing a protein. One of the DNA sequences or its fragment can effectively activate transcription of a gene encoding a target protein irrespective of the presence or absence of Glucose. The DNA sequences and their fragments can constantly activate gene transcription. This polynucleotide sequence represents one of the 74 filamentous fungi Aspergillus genus DNA sequences of the invention.

XX Sequence 538 BP; 128 A; 123 C; 132 G; 155 T; 0 U; 0 Other;

SQ Query Match Score 84.0%; Score 16.8; DB 9; Length 538;  
Best Local Similarity 90.0%; Pred. No. 1e+02;  
Matches 18; Conservative 0; N mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGTTTG 20  
Db 494 CTGGACAGGTAGGGTTTG 475

RESULT 8  
ABZ56309/c  
ID ABZ56309 standard; CDNA; 589 BP.

XX DT 28-MAR-2003 (first entry)

AC ABZ56309;

XX DT 28-MAR-2003 (first entry)

DE Aspergillus oryzae polynucleotide SEQ ID NO 5422.

XX KW Aspergillus oryzae; fermentation; fungus; industrial; EST;

XX expressed sequence tag; gene; ss.

XX OS Aspergillus oryzae.

XX PN WO200279476-A1.

XX PD 10-OCT-2002.

XX FF 22-MAR-2002; 2002WO-IB000890.

XX PR 30-MAR-2001; 2001JP-0008371.

XX PT (NAA/-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX PA (NARE-) NAT RES INST BREWING

XX PA (NORQ) NAT FOOD RES INST MIN AGRIC.

XX PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;

XX PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;

XX DR 2003-046817/04.

PT Detection of expression of specific Aspergillus genes for monitoring the fermentation and growth conditions of the fungus, using DNA probes.  
PS Claim 1; SEQ ID NO 5422; 48pp + Sequence Listing; Japanese.  
XX The invention relates to polynucleotide which are expressed by a fungus under specific culture conditions including one or more of eutrophic, oligotrophic, solid, early germination, alkaline, high temperature, low temperature or maltose culture or Polynucleotides stringently hybridising to these sequences. The polynucleotides are useful for monitoring the progress of fermentation and the growth conditions of a fungus, especially of Aspergillus oryzae which is widely used in industrial fermentation. Also monitoring for fungal contamination. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 589 BP; 143 A; 210 C; 95 G; 141 T; 0 U; 0 Other;

XX Query Match Score 84.0%; Score 16.8; DB 7; Length 589;

XX Best Local Similarity 90.0%; Pred. No. 1e+02;

XX Matches 18; Conservative 0; N mismatches 0;

XX Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGTTTG 20

Db 100 CTGGACAGGTAGGGTTTG 81

RESULT 9  
AAF14551

CC ID AAF14551 standard; CDNA; 809 BP.

XX AC AAF14551;

XX DT 13-MAR-2001 (first entry)

XX XX DE Aspergillus oryzae EST SEQ ID NO:7054.

XX KW Multiple gene expression; filamentous fungal cell; EST;

XX KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;

XX KW Aspergillus oryzae; identification; recombination;

XX KW culture condition; environmental stress; spore morphogenesis;

XX KW metabolic pathway engineering; catabolic pathway engineering; ss.

OS Aspergillus oryzae.

XX PN WO200056762-A2.

XX PD 28-SEP-2000.

XX PR 22-MAR-2000; 2000WO-US007781.

XX PR 22-MAR-1999;

XX PA (NOVO ) NOVO NORDISK BIOTECH INC.

XX PA (NOVO ) NOVO NORDISK AS.

XX PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

XX DR WPI; 2000-594572/56.

XX PR Monitoring differential expression of genes in filamentous fungal cells

XX PT uses fluorescence-labelled nucleic acids isolated from the cells and a

XX substrate of expressed sequence tags.

XX Claim 88; Page 2863-2864; 316pp; English.

XX CC The present invention describes a method for monitoring differential

CC expression of genes in a first filamentous fungal (FF) cell relative to

CC expression of the same genes in one or more second filamentous fungal

CC cells. The method uses fluorescence-labelled nucleic acids isolated from

CC the FF cells and a substrate of expressed sequence tags (EST).

CC The ESTS are used in the methods for monitoring differential expression of genes

CC in a first filamentous fungal (FF) cell relative to expression of the

CC same genes in one or more second filamentous fungal cells. Monitoring the

CC global expression of genes from FF cells allows the production potential

CC of the microorganisms to be improved. New genes may be discovered,

CC possible functions of unknown open reading frames can be identified and

CC gene copy number variation and stability can be monitored. The expression

CC of genes can be used to study how FF cells adapt to changes in culture

CC conditions, environmental stress, spore morphogenesis, recombination,

CC metabolic or catabolic pathway engineering. Using ESTs provides several

CC advantages over genomic or random DNA clones including minimization of

CC redundancy as one spot on an array equals one gene or open reading frame,

CC and organisation of the microarray based on function of the gene.

CC products to facilitate analysis of the results. AAF1478 to AAF1483 represents

CC ESTs from Fusarium venenatum; AAF1484 to AAF14873 represents ESTs from

CC Aspergillus niger; AAF14854 to AAF14873 represents ESTs from

CC Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from  
 CC Trichoderma reesei, which are all specifically claimed in the present  
 invention  
 XX Sequence 809 BP; 1B2 A; 265 C; 161 G; 199 T; 0 U; 2 Other;  
 CC SQ Score 16.8%; DB 3; Length 809;  
 CC Query Match 84.0%; Score 16.8%; DB 3;  
 CC Best Local Similarity 90.0%; Pred. No. 1.1e+02;  
 CC Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC DE Human DAD1 carcinoma associated gene, SEQ ID NO:1478.  
 CC KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;  
 CC prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
 CC KW gene; ds.  
 CC XX Homo sapiens.  
 CC OS XX Homo sapiens.  
 CC PN XX WO2003057146-A2.  
 CC XX XX  
 CC ID XX ADA53227 standard; CDNA; 2305 BP.  
 CC XX PD 17-JUL-2003.  
 CC AC XX ADA53227;  
 CC DT 20-NOV-2003 (first entry)  
 CC XX PR 26-DEC-2002; 2002WO-US041414.  
 CC DB XX PR 26-DEC-2001; 2001US-00035832.  
 CC XX PA (SAGR-) SAGRES DISCOVERY.  
 CC XX PI Morris DW;  
 CC DR XX WPI; 2003-58706/55.  
 CC XX PT New recombinant nucleic acid encoding carcinoma associated protein,  
 CC PR useful for preparing compositions for treating carcinomas.  
 CC XX PA (SAGR-) SAGRES DISCOVERY.  
 CC XX PS Claim 1; SEQ ID NO 1478; 245PP; English.  
 CC XX  
 CC The invention relates to recombinant carcinoma associated (CA) nucleic  
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to  
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The  
 CC invention also encompasses expression vectors and host cells comprising a  
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically  
 CC binds to the protein, and a biochip comprising CA nucleic acid or  
 CC fragments thereof. The sequences of the invention were identified using  
 CC oncogenic retroviruses, which insert into the genome of the host organism  
 CC at random. Many of these do not carry transduced host oncogenes or  
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a  
 CC direct consequence of the effects of proviral integration into host  
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
 CC leukaemia) or a propensity to carcinoma by determination of the sequence  
 CC of a CA gene, or by determination of CA gene expression in particular  
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
 CC therapeutic agents and in screening and evaluating drug candidates. The  
 CC present sequence represents a specifically claimed human CA nucleic acid  
 CC sequence of the invention. Note: The complete sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

RESULT 11  
 ADA02960/C  
 ID ADA02960 standard; DNA; 44325 BP.  
 XX AC ADA02960;  
 XX AX ADA02960;  
 XX DT 06-NOV-2003 (first entry)  
 XX XX  
 XX Human DAD1 carcinoma associated gene, SEQ ID NO:1478.  
 XX DE Human DAD1 carcinoma associated gene, SEQ ID NO:1478.  
 XX KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;  
 XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
 XX KW gene; ds.  
 XX XX Homo sapiens.  
 XX OS XX Homo sapiens.  
 XX PN XX WO2003057146-A2.  
 XX XX  
 XX ID XX ADA53227 standard; CDNA; 2305 BP.  
 XX XX PD 17-JUL-2003.  
 XX AC XX ADA53227;  
 XX DT 20-NOV-2003 (first entry)  
 XX XX PR 26-DEC-2002; 2002WO-US041414.  
 XX DB XX PR 26-DEC-2001; 2001US-00035832.  
 XX XX PA (SAGR-) SAGRES DISCOVERY.  
 XX XX PI Morris DW;  
 XX DR XX WPI; 2003-58706/55.  
 XX XX PT New recombinant nucleic acid encoding carcinoma associated protein,  
 XX PR useful for preparing compositions for treating carcinomas.  
 XX XX PA (SAGR-) SAGRES DISCOVERY.  
 XX XX PS Claim 1; SEQ ID NO 1478; 245PP; English.  
 XX XX  
 CC The invention relates to recombinant carcinoma associated (CA) nucleic  
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to  
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The  
 CC invention also encompasses expression vectors and host cells comprising a  
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically  
 CC binds to the protein, and a biochip comprising CA nucleic acid or  
 CC fragments thereof. The sequences of the invention were identified using  
 CC oncogenic retroviruses, which insert into the genome of the host organism  
 CC at random. Many of these do not carry transduced host oncogenes or  
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a  
 CC direct consequence of the effects of proviral integration into host  
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
 CC leukaemia) or a propensity to carcinoma by determination of the sequence  
 CC of a CA gene, or by determination of CA gene expression in particular  
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
 CC therapeutic agents and in screening and evaluating drug candidates. The  
 CC present sequence represents a specifically claimed human CA nucleic acid  
 CC sequence of the invention. Note: The complete sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

RESULT 12  
 ADB72698/C  
 ID ADB72698 standard; DNA; 44325 BP.  
 XX XX  
 XX Sequence 2305 BP; 457 A; 730 C; 621 G; 497 T; 0 U; 0 Other;  
 CC CC Query Match 84.0%; Score 16.8%; DB 8; Length 44325;  
 CC CC Best Local Similarity 90.0%; Pred. No. 1.5e-02;  
 CC CC Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC CC SQ Sequence 2305 BP; 457 A; 730 C; 621 G; 497 T; 0 U; 0 Other;  
 CC CC The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA54710) and their coding sequences (ADA5233-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 CC XX Sequence 2305 BP; 457 A; 730 C; 621 G; 497 T; 0 U; 0 Other;  
 CC CC Query Match 84.0%; Score 16.8%; DB 7; Length 2305;  
 CC CC Best Local Similarity 90.0%; Pred. No. 1.2e+02;  
 CC CC Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC CC QY 1 CTGGACAGTTAGGGCTTGTG 20  
 CC CC DB 38075 CTGGACAGTTAGGGCTTGTG 38056  
 CC CC  
 CC CC Sequence 809 BP; 11218 A; 9675 C; 9902 G; 13530 T; 0 U; 0 Other;  
 CC CC Query Match 84.0%; Score 16.8%; DB 8; Length 44325;  
 CC CC Best Local Similarity 90.0%; Pred. No. 1.5e-02;  
 CC CC Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC CC QY 1 CTGGACAGTTAGGGCTTGTG 20  
 CC CC DB 38075 CTGGACAGTTAGGGCTTGTG 38056  
 CC CC  
 CC CC Sequence 2118 BP; 457 A; 730 C; 621 G; 497 T; 0 U; 0 Other;  
 CC CC The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA54710) and their coding sequences (ADA5233-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 CC XX Sequence 2118 BP; 457 A; 730 C; 621 G; 497 T; 0 U; 0 Other;  
 CC CC Query Match 84.0%; Score 16.8%; DB 7; Length 2118;  
 CC CC Best Local Similarity 90.0%; Pred. No. 1.2e+02;  
 CC CC Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human DNA (shown in Table 2 of the  
 CC specification) which encodes one of the polypeptides of the invention  
 CC which is differentially expressed during pain. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic form directly from WIPO at  
 CC [ftp://wipo.int/pub/published\\_pct/sequences](ftp://wipo.int/pub/published_pct/sequences).

XX Sequence 2446 BP; 567 A; 637 C; 596 G; 646 T; 0 U; 0 Other;

SQ Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GGACAGGTTAGGGCTTG 20

Db 1850 GGACAGGTTAGGGCTTG 1867

RESULT 15

ADBE57892

ID ADBE57892 standard; DNA; 2446 BP.

AC ADBE57892;

XX DT 29-JUN-2004 (first entry)

XX Human Gene U72649, SEQ ID NO 3758.

XX Human; ds; gene; pain; neuronal tissue; gene therapy; CCI;

XX spinal segmental nerve injury; SNI; Chung.

XX Homo sapiens.

XX OS

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2003; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI ; 2003-268312/26.

DR GENBANK; U72649.

XX

PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human DNA (shown in Table 2 of the  
 CC specification) which encodes one of the polypeptides of the invention  
 CC which is differentially expressed during pain. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic form directly from WIPO at  
 CC [ftp://wipo.int/pub/published\\_pct/sequences](ftp://wipo.int/pub/published_pct/sequences).

XX Sequence 2446 BP; 567 A; 637 C; 596 G; 646 T; 0 U; 0 Other;

SQ Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length 2446;

Best Local Similarity 94.4%; Pred. No. 1.9e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX Query Match 82.0%; Score 16.4; DB 9; Length

Result No.	Score	Query	Match Length	DB ID	Description
C 1	20	100.0	616	13 BX3833606	BX3833606 linear EST 08-MAY-2003
C 2	20	100.0	778	12 BG178487	LOCUS (TURKAT CELL LINE) COT 10-NORMALIZED
C 3	20	100.0	878	12 BQ053486	DEFINITION Homo sapiens cDNA clone CS0DJ13YK10 5'-PRIME, mRNA sequence.
C 4	20	100.0	986	12 BQ054265	ACCESSION BX3833606.1 GI:30457152
					VERSION EST.
					KEYWORDS SOURCE Homo sapiens (human)
					ORGANISM Homo sapiens
					Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
					REMARKS 1 (bases 1 to 616)
					AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
					TITLE Unpublished (2001)
					JOURNAL Unpublished (2001)
					COMMENT Contact: Genoscope - Centre National de Séquençage
					Genoscope - Centre National de Séquençage
					Email: seqref@genoscope.cns.fr
					Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9825.r. For more information about this cluster, see
					http://www.genoscope.cns.fr/
					cgi-bin/cluster.cgi?seq=CS0DJ13YK10&cluster=9825.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
					SUMMARIES
					Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

FEATURES Faraday Avenue Genoscope sequence ID : CSODJ013BF050P1.  
 Source Location/Qualifiers 1..616  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9605"  
 /clone="CSODJ013YK10"  
 /cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
 /cell\_line="JURKAT"  
 /clone\_id="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-Oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 ORIGIN  
 Query Match 100.0%; Score 20; DB 13; Length 616;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CTGGACAGTTAGGGCTTTG 20  
 Db 350 CTGACAGTTAGGGCTTTG 331

RESULT 2  
 BG178487/c  
 LOCUS BG178487 778 bp mRNA linear EST 06-FEB-2001  
 DEFINITION NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4429896 5',  
 mRNA sequence.  
 ACCESSION BG178487  
 VERSION BG178487.1 GI:12685190  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 COMMENT Contact: Robert Straussberg, Ph.D.  
 Email: cgaps-r@mail.nih.gov  
 Tissue Procurement: DCTP/DTP  
 DNA Sequencing by: Incyte Genomics, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Life Technologies, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 http://image.llnl.gov  
 Plate: LLAM1082 row: i column: 01  
 High Quality sequence stop: 657.  
 Location/Qualifiers 1..778  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9605"  
 /clone="IMAGE:4429896"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH MGC 91"  
 /note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: NotI;  
 site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 100.0%; Score 20; DB 12; Length 878;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 FEATURES source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9605"  
 /clone="IMAGE:4429896"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH MGC 91"  
 /note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: NotI;  
 site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 100.0%; Score 20; DB 12; Length 778;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 FEATURES source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9605"  
 /clone="IMAGE:4429896"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH MGC 91"  
 /note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: NotI;  
 site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

RESULT 3  
 BQ053486/c  
 LOCUS BQ053486  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BQ053486  
 VERSION BQ053486.1 GI:19812826  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 COMMENT Contact: Robert Straussberg, Ph.D.  
 Email: cgaps-r@mail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 DNA Sequencing by: Ageacourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 http://image.llnl.gov  
 Plate: LLAM2122 row: 1 column: 06  
 High Quality sequence stop: 394.  
 FEATURES source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:593523"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH MGC 106"  
 /note="Organ: blood; Vector: POTB7; Site 1: XbaI; Site 2:  
 EcorI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcorI/XbaI sites using the following 5' adaptor:  
 GGCACTAG (G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA Synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN  
 Query Match 100.0%; Score 20; DB 12; Length 878;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 FEATURES source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9605"  
 /clone="IMAGE:4429896"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH MGC 106"  
 /note="Organ: prostate; Vector: PCMV-SPORT6; Site 1: NotI;  
 site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

RESULT 4  
 BQ054265/c  
 LOCUS BQ054265  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BQ054265  
 VERSION BQ054265.1 GI:19813605  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 986)

AUTHORS NIH-MGC <http://mgc.ncbi.nih.gov/>  
 TITLE Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps@mail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 CDNA Library Preparation: Rubin Laboratory  
 Consortium (LLNL)  
 DNA Sequencing by: The I.M.A.G.E. Consortium  
 (LLNL)  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2125 row: i column: 12  
 High quality sequence stop: 515.  
 Location/Qualifiers

1. .986  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5936339"  
 /tissue\_type="natural killer cells, cell line"  
 /clone\_host="DH10B (phage-resistant)"  
 /note="Organ: blood; Vector: pOTB7; Site\_1: XbaI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGACGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 1020;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGCTTG 20  
 Db 115 CTGGACAGGTAGGGCTTG 96

FEATURES source

1. .986  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5936339"  
 /tissue\_type="natural killer cells, cell line"  
 /clone\_host="DH10B (phage-resistant)"  
 /note="Organ: blood; Vector: pOTB7; Site\_1: XbaI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGACGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 986;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGCTTG 20  
 Db 182 CTGGACAGGTAGGGCTTG 163

RESULT 5

BQ04281/c BQ054281 1020 bp mRNA linear EST 29-MAR-2002  
 DEFINITION AGENCOURT 6830234 NIH MGC\_106 Homo sapiens cDNA clone IMAGE:5936362  
 5', mRNA sequence.

ACCESSION BQ054281  
 VERSION BQ054281.1 GI:19813621  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Bacteria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)  
 PUBLISHED (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps@mail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 CDNA Library Preparation: Rubin Laboratory  
 Consortium (LLNL)  
 DNA Sequencing by: The I.M.A.G.E. Consortium  
 (LLNL)  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2125 row: j column: 11  
 High quality sequence stop: 556.  
 Location/Qualifiers

1..1020  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"

FEATURES source

/db\_xref="taxon:9606"  
 /clone="IMAGE:5936362"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: blood; Vector: pOTB7; Site\_1: XbaI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGACGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 1020;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGCTTG 20  
 Db 115 CTGGACAGGTAGGGCTTG 96

RESULT 6

AL541041/C AL541041 1201 bp mRNA linear EST 12-MAY-2003  
 DEFINITION Homo sapiens cDNA clone CS0DE005YK23  
 5'-PRIME mRNA sequence.

ACCESSION AL541041  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE Li,W.B., Gruber,C., Jesse,J. and Polayes,D.  
 AUTHORS TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001).  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12871733.  
 CONTACT Genoscope - Centre National de Sequencage  
 Genoscope - Cedex - France  
 BP 191 91005 EVRY Cedex - France  
 Email: seqref@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence Clustter 9825.r. For more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DE005AF12QP1&cluster=9825.r>. Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DE005AF12QP1.  
 FEATURES source  
 1. .1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE005YK23"  
 /tissue\_type="PLACENTA"  
 /note="Vector: pcMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRI sites of the pcMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1201;  
 Best Local Similarity 100.0%; Pred. No. 99;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGCTTG 20

Db 362 CTGGACAGTTAGGCCTTG 343  
 RESULT 7  
 LOCUS BU327205 901 bp mRNA linear EST 28-NOV-2002  
 DEFINITION 603490679\_F1 CSEQCHN63 Gallus gallus cDNA clone ChEST392m20 5', mRNA sequence.  
 ACCESSION BU327205  
 VERSION 1 GI:25835206  
 KEYWORDS EST  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Archosauria; Aves; Neognathae; Craniata; Vertebrata; Euteleostomi;  
 Phasianinae; Gallus.  
 REFERENCE (bases 1 to 901)  
 AUTHORS Boardman, P.E., Sanz-Esguerra, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 TITLE A Comprehensive Collection of Chicken cDNAs  
 JOURNAL Curr. Biol. 12 (22), 1865-1869 (2002)  
 MEDLINE 2145534  
 PUBMED 1245392  
 COMMENT Contact: Simon Hubbard  
 Department of Biomedolecular Sciences  
 University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES source  
 /organism="Gallus gallus"  
 /mol type="mRNA"  
 /strand="White Leghorn, Hisex"  
 /do\_xref="taxon:901"  
 /clone="ChEST392m20"  
 /dev stage="36"  
 /lab host="DH10B"  
 /clone lib="CSEQCHN63"  
 /note="Organ: heads; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an Oligo(dT) primer, using methylated C in the first strand reaction, double-stranded cDNA following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adaptors, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1998) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996) 791, except that a significantly longer reannealing hybridization was used."

ORIGIN Query Match 92.0%; Score 18.4%; DB 13; Length 901;  
 Best Local Similarity 95.0%; Pred. No. 5e+02; Matches 19; Conservative 0; Indels 0; Gaps 0;  
 Qy 1 CTGGACAGTTAGGCCTTG 20  
 Db 659 CTGGACAGTTAGGCCTTG 678

VERSION CA987486.1 GI:27520161  
 KEYWORDS EST  
 SOURCE Xenopus laevis (African clawed frog)  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;  
 Xenopodidae; Xenopus.  
 REFERENCE 1 (bases 1 to 962)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgaps-r@mail.nih.gov  
 Tissue Procurement: Martha Rebbeck, Steven L. Klein, Ph.D.  
 Tissue Preparation: Martha Rebbeck, Steven L. Klein, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: NCI-CGAP Clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at:  
 http://Image.lnl.gov  
 Plate: LILN1484 Row: 9 Column: 24  
 High Quality sequence start: 228  
 High Quality sequence stop: 380.  
 FEATURES source  
 /organism="Xenopus laevis"  
 /mol type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="IMAGE:6864145"  
 /tissue\_type="embryo" (stage 10)  
 /lab\_host="DH10B" (phage-resistant)  
 /clone lib="NICHID\_XGC\_Emb1"  
 /note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI;  
 Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.55 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."  
 ORIGIN Query Match 92.0%; Score 18.4%; DB 14; Length 962;  
 Best Local Similarity 95.0%; Pred. No. 5.1e+02; Matches 19; Conservative 0; Indels 0; Gaps 0;  
 Qy 1 CTGGACAGTTAGGCCTTG 20  
 Db 840 CTGGACAGTTAGGCCTTG 821

RESULT 9  
 LOCUS AA719167//C  
 DEFINITION AA719167//C  
 mRNA sequence.  
 ACCESSION AA719167  
 VERSION AA719167.1 GI:272266  
 EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 332)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgaps-r@mail.nih.gov  
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP Clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at:

RESULT 8  
 LOCUS CA987486/C  
 DEFINITION AGENCOURT\_11284561 NICHD\_XGC\_Emb1 Xenopus laevis cDNA clone IMAGE:6864145 5', mRNA sequence.  
 ACCESSION CA987486

FEATURES	FEATURES
Source	Source
/organism="Homo sapiens"	/organism="Homo sapiens"
/mol_type="mRNA"	/mol_type="mRNA"
/db_Xref="taxon:9606"	/db_Xref="taxon:9606"
/clone="1292678"	/clone="1292678"
/sex="male"	/sex="male"
/lab_host="DH10B"	/lab_host="DH10B"
/clone_lib="Soares testis_NHT"	/clone_lib="Soares testis_NHT"
/note="vector: pRT3D-Pac (Pharmacia) with a modified Polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo (dT) primer [5], TGTATCCGATTCGAACTGGCAACGGGCCATTCTTCTTCTT 3' ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization to Cots5, and was constructed by Bento Soares and M. Fatima Bonaldo."	/note="vector: pRT3D-Pac (Pharmacia) with a modified Polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo (dT) primer [5], TGTATCCGATTCGAACTGGCAACGGGCCATTCTTCTTCTT 3' ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization to Cots5, and was constructed by Bento Soares and M. Fatima Bonaldo."
Location/Qualifiers	Location/Qualifiers
POLYLINKER	POLYLINKER
primer: -40ml3 fwd, ET from Amersham High quality sequence stop: 316.	primer: -40ml3 fwd, ET from Amersham High quality sequence stop: 316.
JOURNAL	JOURNAL
COMMENT	COMMENT
FEATURES	FEATURES
source	source
/organism="Homo sapiens"	/organism="Homo sapiens"
/mol_type="mRNA"	/mol_type="mRNA"
/db_Xref="taxon:9606"	/db_Xref="taxon:9606"
/clone="1292678"	/clone="1292678"
/sex="male"	/sex="male"
/lab_host="DH10B"	/lab_host="DH10B"
/clone_lib="Soares testis_NHT"	/clone_lib="Soares testis_NHT"
/note="vector: pRT3D-Pac (Pharmacia) with a modified Polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo (dT) primer [5], TGTATCCGATTCGAACTGGCAACGGGCCATTCTTCTTCTT 3' ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization to Cots5, and was constructed by Bento Soares and M. Fatima Bonaldo."	/note="vector: pRT3D-Pac (Pharmacia) with a modified Polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo (dT) primer [5], TGTATCCGATTCGAACTGGCAACGGGCCATTCTTCTTCTT 3' ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization to Cots5, and was constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN	ORIGIN
Query Match Score 87.0%; Best Local Similarity 94.7%; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Query Match Score 87.0%; Best Local Similarity 94.7%; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGGACAGGTAGGGCTT 19	Qy 1 CTGGACAGGTAGGGCTT 19
Db 207 CTGGACAGGTAGGGCTT 189	Db 207 CTGGACAGGTAGGGCTT 189
RESULT 11	RESULT 11
LOCUS BX089099	LOCUS BX089099
DEFINITION Soares testis_NHT	DEFINITION Soares testis_NHT
ACCESION BX089099	ACCESION BX089099
VERSION BX089099.1	VERSION BX089099.1
KEYWORDS EST.	KEYWORDS EST.
SOURCE Homo sapiens	SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 413)	REFERENCE 1 (bases 1 to 413)
AUTHORS Ebert,L.; Heil,O.; Henning,S.; Neubert,P.; Partsch,E.; Peters,M.; Radloff,U.; Schnedl,D. and Korn,B.	AUTHORS Ebert,L.; Heil,O.; Henning,S.; Neubert,P.; Partsch,E.; Peters,M.; Radloff,U.; Schnedl,D. and Korn,B.
TITLE Human UnigeneSet - RZPD3	TITLE Human UnigeneSet - RZPD3
JOURNAL Unpublished (2003)	JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs	COMMENT Contact: Ina Rolfs
ACCESSION RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH	ACCESSION RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
DEFINITION Im Neuenheimer Feld 500, D-69120 Heidelberg, Germany	DEFINITION Im Neuenheimer Feld 500, D-69120 Heidelberg, Germany
LOCUS RZPDlib_I.M.A.G.E. CDNA Clone Collection; RZPDlib_N.972	LOCUS RZPDlib_I.M.A.G.E. CDNA Clone Collection; RZPDlib_N.972
ORGANISM Human UnigeneSet - RZPDs (RZPDlib_N.972)	ORGANISM Human UnigeneSet - RZPDs (RZPDlib_N.972)
REFERENCE http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972	REFERENCE http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972
AUTHORS RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH	AUTHORS RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
TITLE Heubnerweg 6, D-14059 Berlin, Germany	TITLE Heubnerweg 6, D-14059 Berlin, Germany
KEYWORDS EST.	KEYWORDS EST.
SOURCE Homo sapiens (human)	SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 357)	REFERENCE 1 (bases 1 to 357)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)	JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.	COMMENT Contact: Robert Strausberg, Ph.D.
ACCESSION E0707117	ACCESSION E0707117
DEFINITION CGAPs -@mail.nih.gov	DEFINITION CGAPs -@mail.nih.gov
ORGANISM CGNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.	ORGANISM CGNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
COMMENT DNA Sequencing by: Greg Lennon, Ph.D.	COMMENT DNA Sequencing by: Greg Lennon, Ph.D.
ACCESSION M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.	ACCESSION M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
DEFINITION NCI-NCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINN at:	DEFINITION NCI-NCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINN at:
JOURNAL www-bio.llnl.gov/bbrp/image/image.html	JOURNAL www-bio.llnl.gov/bbrp/image/image.html
COMMENT Insert Length: 466 Std Error: 0.00	COMMENT Insert Length: 466 Std Error: 0.00
ACCESSION M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.	ACCESSION M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
DEFINITION Location/Qualifiers	DEFINITION Location/Qualifiers
source	source
/organism="Homo sapiens"	/organism="Homo sapiens"
/mol_type="mRNA"	/mol_type="mRNA"
/db_Xref="taxon:9606"	/db_Xref="taxon:9606"
/clone="1292678"	/clone="1292678"
/sex="male"	/sex="male"
/lab_host="DH10B"	/lab_host="DH10B"
/clone_lib="Soares testis_NHT"	/clone_lib="Soares testis_NHT"
/note="vector: pRT3D-Pac (Pharmacia) with a modified Polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo (dT) primer [5], TGTATCCGATTCGAACTGGCAACGGGCCATTCTTCTTCTT 3' ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization to Cots5, and was constructed by Bento Soares and M. Fatima Bonaldo."	/note="vector: pRT3D-Pac (Pharmacia) with a modified Polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo (dT) primer [5], TGTATCCGATTCGAACTGGCAACGGGCCATTCTTCTTCTT 3' ]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization to Cots5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN constructed by Bento Soares and M. Fatima Bonaldo. "

Query Match Similarity 87.0%; Score 17.4%; DB 13; Length 413; Best Local Similarity 94.7%; Pred. No. 1.3e+03; Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTAGGGCTTT 19  
Db 214 CTGGACAGGTAGGGCTTT 232

RESULT 12 AG048198/C AG048198 Pan troglodytes DNA, clone: PTB-027P05.R, genomic survey sequence.

DEFINITION Pan troglodytes DNA, clone: PTB-027P05.R, genomic survey sequence.

ACCESSION AG048198  
VERSION AG048198.1  
KEYWORDS GSS.  
ORGANISM Pan troglodytes (chimpanzee)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1. Fuijyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tocoki,Y., Watanabe,H. and Sakaki,Y.  
TITLE BAC end sequences of Library PTB  
JOURNAL Unpublished

2 (bases 1 to 740)  
Fuijyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tocoki,Y., Watanabe,H. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimmpdb@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)  
COMMENT Clones are derived from the Chimpzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS Sequencing: M13Rev

LIBRARY Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI

FEATURES source  
LOCATIONS Qualifiers 1..740  
1.740  
1.740

ORGANISM "Pan troglodytes"  
/mol type="genomic DNA"  
/db\_xref="ITaxon:9558"  
/clone="PTB-027P05.R"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"

Query Match Similarity 87.0%; Score 17.4%; DB 29; Length 740; Best Local Similarity 94.7%; Pred. No. 1.4e+03; Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TGGACAGGTAGGGCTTG 20  
Db 583 TGGACAGGTAGGGCTTG 565

KEYWORDS SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus .  
REFERENCE 1 (bases 1 to 177)  
AUTHORS Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J., Piggett,J., BeltranndeRio,H., Buxton,E.C., Edwards,J., Finch,R.A., Fiddle,C.J., Guppa,A., Hansen,G., Hu,Y., Huang,W., Jaing,C., Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markevich,D., Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrock,J., Shi,Z.-Z., Sparks,M.J., Van Sligghorst,I., Vogel,P., Walke,W., Xu,N., Zhu,Q., Person,C. and Sands,A.T.

TITLE JOURNAL COMMENT screen to identify potential targets for therapeutic intervention Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
Contact: Zambrowicz,B.P.  
OmniBank  
Lexicon Genetics Incorporated  
4000 Research Forest Drive, The Woodlands, TX 77381, USA  
Email: materials@lexgen.com  
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392 (6676):608-11)  
Class: Gene Trap  
Location/Qualifiers 1..177  
/organism="Mus musculus"  
/mol type="genomic DNA"  
/strain="129sv/Ev"  
/db\_xref="ITaxon:10090"  
/clone="OST246195"  
/cell\_type="embryonic stem cell"  
/clone\_lib="Mus musculus 129sv/Ev"

ORGANISM ORIGIN Query Match Similarity 84.0%; Score 16.8%; DB 29; Length 177; Best Local Similarity 90.0%; Pred. No. 2.1e+03; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CTGGACAGGTAGGGCTTG 20  
Db 44 CTGGACAGGTAGGGCTTG 25

LIBRARY RESULT 14  
Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI  
FEATURES source  
LOCUS CB487227/c  
DEFINITION omkrtit001012 Oncorhynchus mykiss reproductive Oncorhynchus mykiss mRNA linear EST 01-APR-2003  
ACCESSION CB487227  
VERSION CB487227.1  
KEYWORDS SOURCE CB487227  
ORGANISM Oncorhynchus mykiss (rainbow trout)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleosteii; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
COMMENT A survey of Salmo salar transcripts from high complexity cDNA  
Reference GRASP Consortium, Davidson,W.S., Koop,B.F. and http://web.uvic.ca/bcr/grasp/  
Authors TITLE JOURNAL COMMENT Contact: Koop,B.F.  
Centre for Biomedical Research  
University of Victoria, Victoria BC, V8W 3N5, Canada  
PO Box 3030 STN CSC, Victoria BC, V8W 3N5, Canada  
Tel: 250 472 4067  
Fax: 250 472 4075  
Email: bkoop@uvic.ca  
Centre for Biomedical Research, University of Victoria cDNA preparation and sequencing: Robert Alberto Marianne Beetz-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, BF Koop.

RESULT 13 CG91538/C LOCUS CG91538  
DEFINITION CGT246195 Mus musculus 129Sv/Ev Mus musculus genomic clone  
ACCESSION CG91538  
VERSION CG91538.1 GI:3739954